# ERATOSTHENIAN AVERAGES

BY

AURAL WINTNER

BALTIMORE

1943

# ERATOSTHENIAN AVERAGES

BY

AURELWINTNER

BALTIMORE

1943

Composed and Printed at the
WAVERLY PRESS, INC.
Baltimore, U. S. A.

# PREFACE

> Je ne doute pas que, pour peu qu'on regarde la progression de ces nombres, on ne désespere presque d'y découvrir le moindre ordre, vu que l'irrégularité de la suite des nombres premiers s'y trouve entremêlée tellement, qu'il semblera d'abord impossible d'indiquer quelque loi ces nombres observent entre eux, sans qu'on sache celle des nombres premiers.
>
> L. Euler, Opera Omnia, ser. 1, vol. 2, pp. 244–245.

The course of lectures on which this monograph is based has centered about two aspects of the distribution of primes; aspects which are closely interrelated and can roughly be indicated by the headings Statistics and Almost-periodicity.

Both of these aspects must be dismissed as mere "physics", if it is assumed that the sole purpose of the analytic theory of numbers is to *abschätzen*. But *Abschätzungen* in themselves can never lead to any insight into the system of primes. For instance, about the only thing that can be prophesied from the upper estimates obtained thus far in Riemann's hypothesis or in Dirichlet's divisor problem is that no solution will ever be supplied by this technique. Actually, there is a variety of more modest problems of substantially arithmetical interest and still of "physical" nature. Some of them turn out to be equivalent to, some others lie slightly deeper than, while still others, though fascinating enough, do not depend on, the prime number theorem. As illustrated by Brun's form of the sieve of Eratosthenes, there are no absolute standards in terms of which it is meaningful to state that results of the latter kind are by necessity of less depth than those depending on the prime number theorem.

The first of the two "physical" aspects referred to at the beginning will appear on all the various levels of the relative depth just mentioned. In fact, that chaotic homogeneity which makes any statistics possible at all is, in its various forms of refinement, a principal characteristic of the asymptotic distribution of the primes. The ergodic behavior in question cannot easily be specified in a sufficiently inclusive form; mainly because the underlying probabilities, being asymptotic probabilities, are defined in terms of relative measures, in contradistinction to the additive measures of the strictly Lebesgueian theory of absolute probabilities.

The other aspect deals with the harmonic analysis of deviations from the underlying average state. In this regard, there will be developed a Fourier theory for "arbitrary" sequences of numbers; a theory which at first glance does not appear to involve the distribution of the primes. Actually, the latter will always lurk in the background of the resulting criteria for almost-periodicity.

To quote Plotinus in a context which he could not fully have appreciated, *αἱ δὲ ἁρμονίαι αἱ ἐν ταῖς φωναῖς ἀφανεῖς τὰς φανερὰς ποιήσασαι* (Enneads, I, vi, 3). It is only another manifestation of this situation that Riemann's hypothesis is equivalent to the assumption that there exists an harmonic analysis for the remainder term of the prime number theorem; cf. the papers referred to under [91] in the Bibliography.

I wish to thank Professor E. K. Haviland for his devoted help in seeing the manuscript through the press, and the Waverly Press for their understanding cooperation.

Oakland, Md., August 1942 AUREL WINTNER

## CONTENTS

### Part I

### ERATOSTHENIAN SUMMATIONS



### Part II

### ARITHMETICAL ALMOST PERIODICITIES



### Part III

### THE STATISTICS OF THE PRIME NUMBER THEOREM

## Part I

## ERATOSTHENIAN SUMMATIONS



### Mean Values

**1.** Unless stated otherwise, let

$$\sum = \sum_{1}^{\infty}. \tag{1}$$

Let $n$, $m$, $k$ denote variable positive integers. Correspondingly, by a function $g = g(n)$ of $n$ will be meant a sequence of numbers $g(1), g(2), \cdots$.

For every function $g(n)$, let a "norm" be defined by

$$N(g) = \limsup_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} | g(m) | , \tag{2}$$

where $N(g) = \infty$ is allowed. On the other hand, let the "mean"

$$M(g) = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} g(m) \tag{3}$$

be considered as undefined unless it exists as a *finite* limit.

As pointed out by Kronecker [56],

$$M(g) = 0 \quad \text{if} \quad \sum g(n)/n \text{ converges} \tag{4}$$

(but not conversely). In fact, if $a_n = g(n)/n$, the assertion (4) is that the convergence of the series $\Sigma\, a_n$ implies that $a_1 + 2a_2 + \cdots + na_n = o(n)$. But the truth of this assertion follows by observing that the ratio of $a_1 + 2a_2 + \cdots + na_n$ to $n$ is identical with the deviation of the $n$-th partial sum of $\Sigma\, a_n$ from the $n$-th $(C, 1)$-sum of $\Sigma\, a_n$ .

It is also seen that

$$g(n) = o(n) \quad \text{if} \quad M(g) \text{ exists} \tag{5}$$

(but not conversely). In fact, (3) is equivalent to

$$\sum_{m=1}^{n} g(m) = nM(g) + o(n). \tag{6}$$

But if $n$ is replaced by $n - 1$ in (6), then (5) follows by subtraction.

It is clear from (6) and (5) that, if $M(g)$ exists, both

$$\sum_{m\,<\,\theta x} g(m) = \theta x M(g) + o(x) \quad \text{and} \quad \sum_{m \leqq \theta x} g(m) = \theta x M(g) + o(x) \quad \text{as} \quad x \to \infty$$

hold not only for $\theta = 1$ but for every fixed $\theta$ (where $0 \leqq \theta \leqq 1$). This means that the mean of $g(n)$ exists if and only if there exists a number, $M(g)$, satisfying

$$\frac{1}{n}\sum_{m=1}^{n} F\left(\frac{m}{n}\right) g(m) \to M(g) \int_0^1 F(t)\,dt \quad \text{as} \quad n \to \infty \tag{7}$$

for the function $F(t)$, $0 \leqq t \leqq 1$, that is 0 or 1 according as $t$ is not or is between 0 and $\theta$, where $t = \theta$ is arbitrarily fixed and both choices $F(\theta) = 0$, $F(\theta) = 1$ are allowed. It follows that if $M(g)$ exists, then (7) holds for every step function $F$ (having a finite number of jumps). In fact, it is clear that if (7) holds for $F = F_1$ and for $F = F_2$, then it holds for $F = c_1F_1 + c_2F_2$ also, where $c_1$, $c_2$ are arbitrary constants.

**2.** In order to include in (7) functions $F$ that are not step functions, it is necessary to subject the function $g(n)$ to suitable Tauberian restrictions of the type considered by Axer [1]. The simplest restriction of this type is the assumption that $g(n)$ has a finite norm (2). In fact, it is clear that if $N(g) < \infty$, and if (7) holds for $F = F_1, F_2, \cdots$, where $\{F_k(t)\}$ is a uniformly convergent sequence of bounded functions, then (7) holds for $F = \lim F_k$ also. Hence, if $N(g) < \infty$, and if $M(g)$ exists, then (7) holds for every $F$ that is uniform limit of step functions; for instance, for every $F$ that either is continuous or has a finite number of discontinuities of the first kind.

The latter condition is not satisfied by*

$$F(t) = t^{-1} - [t^{-1}], \quad (0 < t \leqq 1;\, F(0) = 0, \text{ say}), \tag{8}$$

an $R$-integrable function (of unbounded variation) which underlies the divisor problem of Dirichlet ([18]; cf. [19]). However, the condition is obviously satisfied by each of the functions $F^1, F^2, \cdots$, if $F^k$ denotes the function attaining the value 0 or $F$ according as $0 \leqq t < k^{-1}$ or $k^{-1} \leqq t \leqq 1$, where $F$ is the function (8). Furthermore, $0 \leqq F^k(t) \leqq F(t) \leqq 1$ for every $t$ and so, if $\epsilon > 0$,

$$\frac{1}{n}\sum_{m/n<\epsilon} F^k\left(\frac{m}{n}\right) g(m) \leqq \frac{1}{n}\sum_{m<n\epsilon} |g(m)| \leqq \epsilon\, \underset{1\leqq x<\infty}{\text{l.u.b.}}\, \frac{1}{x}\sum_{m<x} |g(m)|.$$

But the least upper bound multiplying $\epsilon$ is independent of $k$, $n$, $\epsilon$, and is finite, since $N(g) < \infty$ by assumption. It follows therefore by letting $k \to \infty$ and $\epsilon \to 0$ (in this order), that (7), being true for every $F^k$, must hold for the limit function (8) also.

It follows that, if $M(g)$ exists and $N(g) < \infty$, then

$$\frac{1}{n}\sum_{m=1}^{n}\left\{\frac{n}{m} - \left[\frac{n}{m}\right]\right\} g(m) \to \Gamma'(2)M(g) \quad \text{as} \quad n \to \infty. \tag{9}$$

In fact, substitution of (8) into (7) gives (9), since the integral on the right of (7) becomes $1 - C$, where $C$ is the Euler-Mascheroni constant; so that $1 - C = 1 + \Gamma'(1) = \Gamma'(2)$.

* The bracket refers to the greatest integer.

The above restriction, $N(g) < \infty$, is more inclusive than the standard Tauberian condition, $g(n) = O_L(1)$, since the case of unilateral boundedness can be reduced to the case $N(g) < \infty$ (but not conversely). In fact, since (9) is reduced to the definition of $\Gamma'(2)$ when $g(n) = 1$ for every $n$, it is clear that if (9) holds for a function $g$ of $n$, then it holds for the function $g +$ const. also. Hence, $g(n) = O_L(1)$ can be reduced to $g(n) \geqq 0$. But then the existence of $M(g)$ implies that $N(g) < \infty$; cf. (2) and (3).

**3.** If $F(t)$, $0 \leqq t \leqq 1$, is an $L$-integrable function, it may be said to possess a density function, say $G(t)$, $0 \leqq t \leqq 1$, if

$$\int_t^{t+h} \{F(u) - G(t)\}\, du = o(|h|) \quad \text{as} \quad h \to 0 \tag{10}$$

holds for every $t$ (according to Lebesgue, not only (10) but also the corresponding relation for the absolute value of the integrand is always true for almost all $t$). The notion which parallels the density function in case $F(t)$ is replaced by a function $f(n)$ of the positive integer $n$ goes back to §301–§304 of the *Disquisitiones Arithmeticae* [26]. In fact, the considerations of Gauss lead to the idea of a density function, say $g(n)$, to be defined (if it exists) as follows:

$$\sum_{l=n+1}^{n+m} \{f(l) - g(n)\} = o(m) \quad \text{as} \quad m \to \infty \tag{11}$$

holds whenever $m/n \to 0$, where $n = n(m)$ is *any* fixed function of $m$. Needless to say, $g(n)$ is not uniquely determined by $f(n)$, since $g(n)$ can be replaced by $g(n) + h(n)$ whenever $h(n) \to 0$ as $n \to \infty$.

In order to make possible a direct comparison of this notion with the idea of a mean $M(f)$ of $f(n)$, only the case will be considered in which $g(n)$ is (or, more precisely, can be chosen as) independent of $n$. Let then the constant $g(n) = g(1)$ be denoted by $G(f)$ and called the Gaussian density of $f(n)$. Thus, if

$$s(n, m) = f(n+1) + \cdots + f(n+m), \tag{12}$$

it is seen from (11) that the existence of a Gaussian density for $f(n)$ means the existence of a number $G(f)$ having the property that

$$s(n, m)/m \to G(f) \quad \text{as} \quad m \to \infty \tag{13}$$

holds for any function $n = n(m)$ satisfying $m = o(n)$ as $m \to \infty$.

The mutual relationship of $M(f)$ and $G(f)$ is often misunderstood in the literature. Bachmann's comments in his *Analytische Zahlentheorie* (pp. 399–400) are misleading, to say the least. The corresponding statements in Kronecker's *Vorlesungen* (pp. 314–372) are provably false, as will be seen at the end of §4 below. It must, however, be said that the careful statements in Kronecker's original paper [56] were quite correct, and so the mistake might have been introduced by Hensel, the editor of the lectures.

**4.** The actual situation is as follows: If $f(n)$ possesses a Gaussian density, then

$$\text{(i)} \qquad f(n) = O(1) \quad \text{and} \quad \text{(ii)} \qquad G(f) = M(f).$$

In (ii), the existence of $M(f)$ is part of the statement, the only assumption being the existence of $G(f)$. On the other hand, obvious examples show that (i) and the existence of $M(f)$ together do not imply the existence of $G(f)$. In this connection, cf. Hartman and Wintner [43].

In view of the definition of $G(f)$, as given by and after (13), the assumption is that, if $m_1 < m_2 < \cdots$ and $n_1 < n_2 < \cdots$, then

$$(14) \qquad s(n_k, m_k)/m_k \to G(f) \quad \text{as} \quad k \to \infty \quad \text{whenever} \quad m_k/n_k \to 0.$$

Let the sequence $m_1', m_2', \cdots$ be defined, in terms of a given sequence $m_1, m_2, \cdots$, by the recursion formula $m_k' = m_{k-1}' + m_k$, where $m_0' = 0$. Suppose that the sequence $m_1, m_2, \cdots$ is so chosen that $m_k/m_k' \to 0$ as $k \to \infty$. Then (14) implies that $s(m_{k-1}', m_k)/m_k \to G(f)$. It follows therefore from Cauchy's lemma on averages, that the ratio of $s(m_0', m_1) + \cdots + s(m_{k-1}', m_k)$ to $m_1 + \cdots + m_k$ also tends to $G(f)$. Since the denominator and the numerator of this ratio are respectively identical with $m_k'$ and with the sum (12) belonging to $n = 0$, $m = m_k'$, it follows that $s(0, m_k')/m_k' \to G(f)$ as $k \to \infty$. On the other hand, it is seen from (12) and (3) that the existence of $M(f)$ means that $s(0, m)/m \to M(f)$ holds as $m$ runs through all positive integers. Since there are $m_k$ positive integers $m$ between $m = m_{k-1}' + 1$ and $m = m_k'$, and since $m_k/m_k' \to 0$ as $k \to \infty$, it now follows easily that, in order to prove (ii), it is sufficient to prove (i). But it is readily seen from (12) that if (i) is false, then it is possible to construct two sequences $n_1 < n_2 < \cdots$, $m_1 < m_2 < \cdots$ which satisfy $m_k/n_k \to 0$ as $k \to \infty$ but are such that the limit of $s(n_k, m_k)/m_k$ as $k \to \infty$ does not exist. However, this contradicts (14).

As an illustration, let $f(n)$ be 1 or 0 according as $n$ is or is not a prime. Then $M(f)$ exists and vanishes, since the number of primes not exceeding $m$ is $s(0, m) = o(m)$. On the other hand, not even the prime number theorem, $s(1, m) \sim m/\log m$ (or, for that matter, its remainder term supplied by Riemann's hypothesis) assures the existence of $G(f)$. That $G(f)$ happens to exist, follows from Brun's form of the sieve of Eratosthenes ([5]; cf. Hardy and Littlewood [39], p. 69). In fact, Brun's elementary approach assures the existence of a constant $C$ satisfying $s(n, m) < Cm/\log(m + 1)$. But all that is needed for the existence (and the vanishing) of $G(f)$ is that $s(n, m) = o(m)$ should hold uniformly in $n$.

Suppose, however, that the preceding $f(n)$ is replaced by the function which is $\log n$ or 0 according as $n$ is or is not a prime. Then the prime number theorem states that $M(f)$ exists and has the value 1. It follows therefore from (ii) that $G(f) = 1$, if $G(f)$ exists. But the number of integers $n$ satisfying $f(n) \neq 0$ and $n \leqq m$ is the preceding $s(0, m) = o(m)$, which implies that $G(f) = 0$, if $G(f)$ exists. Actually, the necessary condition (i) is violated.

## The Eratosthenian Matrix

**5.** If either of two arbitrary functions, say $f = f(n)$ and $f' = f'(n)$, of the positive integer $n$ is given, the other function is uniquely determined by the assignment

$$f(n) = \sum_{d \mid n} f'(d), \qquad (n = 1, 2, \cdots) \tag{15}$$

(the summation index, $d$, runs through all divisors of $n$). In fact, if (15) is thought of as a linear transformation

$$f(n) = \sum_{m=1}^{\infty} \epsilon_{nm} f'(m), \tag{16}$$

then those elements of the infinite square-matrix $(\epsilon_{nm})$ which are situated above its principal diagonal all vanish, since $d \mid n$ is impossible when $d > n$, and so there exists a unique inverse matrix $(\epsilon_{nm})^{-1}$, since $\epsilon_{nn} \neq 0$ in view of $n \mid n$.

It is clear from (15) that, if $p$ is a prime, then

$$f(p^k) = \sum_{j=0}^{k} f'(p^j), \quad \text{hence} \quad f'(p^k) = f(p^k) - f(p^{k-1}), \tag{17}$$

where $k = 1, 2, \cdots$ and

$$f(1) = f'(1). \tag{18}$$

However, the values of the function can be assigned arbitrarily for every $n$, and (17), (18) do not determine the function values for any $n$ divisible by more than one prime.

Since (16) is identical with (15), every $\epsilon_{nm}$ is either 0 or 1, i.e., of the form $\frac{1}{2} \pm \frac{1}{2}$, where the sign is a single-valued function of $n$ and $m$ together. According to (15), the actual determination of this function appears to require the knowledge of all the divisibility properties of all the integers. Fortunately, this monstrous task can be avoided by approaching the matrix $(\epsilon_{nm})$ vertically, and not, as in (15), horizontally. In fact, the function $\epsilon_{nm} = \frac{1}{2} \pm \frac{1}{2}$ can be described as follows: The elements of the $m$-th column of the infinite square-matrix $(\epsilon_{nm})$ form that periodic sequence, of period $m$, in which $m - 1$ zeros are followed by a single 1. Correspondingly, (15) is equivalent to

$$\sum_{m=1}^{n} f(m) = \sum_{m=1}^{\infty} \left[\frac{n}{m}\right] f'(m) \equiv \sum_{m=1}^{n} \left[\frac{n}{m}\right] f'(m), \tag{19}$$

where $[x]$ denotes the greatest integer not exceeding $x$.

The actual inventor of the linear transformation (15), that is, of the matrix $(\epsilon_{nm})$, was neither Möbius [64], nor Dedekind [15] and Liouville [62], but Eratosthenes [77], [65] (needless to say, none of them used the terminology of linear transformations or matrices). In fact, the possibility of constructing the matrix

in terms of its periodic columns is a mere restatement of the proof of the sieve of Eratosthenes.* An equivalent description of the periodic structure is supplied by the sequences which de Polignac [67] calls diatomic (δια τέμνω). As pointed out recently ([89], p. 147), even the group of the (bounded) non-singular $D$-matrices of Toeplitz [80] can be thought of as a representation of the sieve process.

The Eratosthenian way of reading (15) corresponds to the replacement of the $R$-integral by the $L$-integral and is therefore indispensable in questions relating to the distribution of the values $f$ attained by the function. The simplest facts in this direction depend on the inequality

$$(20) \qquad \left| \sum_{m=1}^{n} f(m) - n \sum_{m=1}^{n} f'(m)/m \right| \leqq \sum_{m=1}^{n} | f'(m) |,$$

which is obvious from (19), if $| [x] - x | < 1$ is applied to $x = n/m$.

It is easy to extend (19) to multilinear forms. For instance, if $\{n, m\}$ denotes the least common multiple of $n$ and $m$, and $\bar{a}$ the complex conjugate of $a$, then†

$$(21) \qquad \sum_{m=1}^{n} | f(m) |^2 = \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} \left[ \frac{n}{\{m, l\}} \right] f'(m)\bar{f}'(l) \equiv \sum \sum_{\{m, l\} \leqq n}$$

(since $[\ ] = 0$ when $\{m, l\} > n$). In fact, it is seen from (15) that the simple sum on the left of (21) is identical with

$$\sum_{d | n} \sum_{c | n} f'(d)\bar{f}'(c), \quad \text{i.e.} \quad \sum_{d=1}^{\infty} \sum_{c=1}^{\infty} \pi_d^c(n) f'(d)\bar{f}'(c),$$

* The construction of the Eratosthenian matrix $(\epsilon_{nm})$ is independent both of the uniqueness of the factorization of $n$ into prime powers and of the existence of an infinity of primes. Hence, both of these theorems of Euclid must be provable on the basis of a sole "graph". In fact, all the divisibility properties of all the integers are expressible in terms of the graph consisting of the lower right quarter of an infinite planar lattice in which every $m$-th, and no other, rectangle of the $m$-th column is marked by an asterisk. For instance, $n$ is a prime if and only if the $n$-th row of the graph contains exactly two asterisks. This *particular* application of the graph is the *theorem* of Eratosthenes; his *proof*, which is not given by Nichomachos [77], [65], depends on the *whole* of the periodic structure.

† If $f'(n) = 0$ for every $n$ not contained in a sequence of integers $k_1, k_2, \cdots$ which are pairwise relatively prime, then, since $\{m, m\} = m$, and since $\{m, l\} = ml$ when $m$ and $l$ are relatively prime, the identity (21) is reduced to

$$\sum_{m=1}^{n} f(m)^2 = \sum_{m=1}^{n} \left[ \frac{n}{m} \right] f'(m)^2 + 2 \sum_{ml \leqq n}^{m<l} \left[ \frac{n}{ml} \right] f'(m)f'(l),$$

if $f$ is real-valued. A revealing application of the latter relation was given by Turán [82]. In his case, $k_1, k_2, \cdots$ is the sequence of all primes.

It is easy to extend Turán's results to the case of an arbitrary sequence $k_1, k_2, \cdots$ satisfying $(k_i, k_j) = 1$ for $i \neq j$. A corresponding remark holds for the results on additive functions, referred to in §67 below.

where $\pi_d^c(n)$ denotes, for a given pair of positive integers $d$, $c$, the number of those positive integers not exceeding $n$ that are divisible by $d$ *and* by $c$, that is, by $\{d, c\}$. Since $\pi_d^c(n)$ is the integral part of the quotient $n/\{d, c\}$, the identity (21) follows.

**6.** Let $\mu(n)$ denote the $n$-th element in the first column of the matrix $(\epsilon_{nm})^{-1}$, that is, of the matrix inverse to the matrix of (16). This means that, in virtue of the one-to-one correspondence (15) between $f$ and $f'$, the function $\mu$ is defined as follows:

$$f' = \mu \quad \text{if} \quad f(1) = 1 \quad \text{and} \quad f(n) = 0 \quad \text{for} \quad n \neq 1. \tag{22}$$

More generally, if $m$ is fixed, and if the function of $n$ representing the $n$-th element in the $m$-th column of $(\epsilon_{nm})^{-1}$ is called $f'(n)$, then the corresponding function $f(n)$ is 0 or 1 according as $n \neq m$ or $n = m$. Hence it is clear from the periodic structure of $(\epsilon_{nm})$ that $(\epsilon_{nm})^{-1}$ is the matrix in which the $n$-th 1 in the $m$-th column of $(\epsilon_{nm})$ is replaced by $\mu(n)$; it being understood that the zeros of $(\epsilon_{nm})$ are zeros of $(\epsilon_{nm})^{-1}$ also. In other words, the inverse matrix can be represented in terms of the elements, $\mu(n)$, of its first column, as follows:

$$(\epsilon_{nm})^{-1} = (\mu(n/m)), \quad \text{where} \quad \mu(x) = 0 \quad \text{unless} \quad x = [x] \tag{23}$$

(that is, $\mu(n/m) = 0$ unless $m \mid n$, where / is a fraction sign).

Since (15) is identical with (16), it follows from (23) that the inverse of the linear substitution (15) is

$$f'(n) = \sum_{d \mid n} \mu(n/d) f(d), \qquad (n = 1, 2, \cdots). \tag{24}$$

In particular, from (22),

$$1 = \mu(1) \quad \text{and} \quad 0 = \sum_{d \mid n} \mu(d) \quad \text{if} \quad n > 1, \tag{24 bis}$$

which means that the function $\mu(n)$ defined by (22) is the function $\mu(n)$ of Möbius. Correspondingly, the inversion of (15) is what Sylvester [74] called a theorem in logic (viz., an identity in Boolean algebra).

Another way of identifying (22) with the explicit definition of $\mu(n)$ follows by writing the term $f'(d)$ of the sum (15) in the form $1.f'(d)$, and observing then that, since $\Sigma\, 1.n^{-s} = \zeta(s)$, the definition (15) is formally equivalent to

$$\sum f(n) n^{-s} = \zeta(s) \sum f'(n) n^{-s}; \tag{25}$$

cf. (1). However, for reasons which will become apparent in §84, Dirichlet series will hardly be used in the sequel.

It should be mentioned for later reference that the function usually denoted by $\Lambda(n)$, that is, the coefficient of $n^{-s}$ in the Dirichlet series of the logarithmic derivative of $1/\zeta(s)$, can, by (15), be defined as follows:

$$f'(n) = \Lambda(n) \quad \text{if} \quad f(n) = \log n. \tag{26}$$

**7.** It is easily verified from (15) that (19) can be generalized to

$$\sum_{m=1}^{n} f(m) = \sum_{m=1}^{l} [n/m]f'(m) + \sum_{m=1}^{n/l} F'(n/m) - [n/l]F'(l), \tag{27}$$

where $l$ is any positive integer not exceeding $n$, the signs / are fraction signs, the brackets refer to greatest integers (so that the second summation on the right is to be arrested at $m = [n/l]$), finally $F'(n)$ denotes the sum function

$$F'(n) = \sum_{m=1}^{n} f'(m) \tag{28}$$

(it being understood that $F'(x) = F'([x])$ when $x$ is not an integer). The identity (27), which goes back to Dirichlet [18], [19], leads to easy and unified proofs, and at the same time to generalizations, of a whole class of classical estimates.

In order to see this, suppose first that $f'(n) = O(1)$, and choose $l = [n^{\frac{1}{2}}]$ in (27). Then omission of the brackets in the first sum on the right of (27) introduces the error $[n^{\frac{1}{2}}]O(1) = O(n^{\frac{1}{2}})$. Hence

$$\sum_{m=1}^{n} f(m) = n\sum_{m=1}^{n^{\frac{1}{2}}} f'(m)/m + \sum_{m=1}^{n^{\frac{1}{2}}} F'(n/m) - n^{\frac{1}{2}}F'(n^{\frac{1}{2}}) + O(n^{\frac{1}{2}}). \tag{$29_1$}$$

For instance, if $f'(n) = 1$ for every $n$, then $f(n)$ is, by (15), the number, $d(n)$, of all divisors of $n$, and (28) is reduced to $F'(n) = n$. Hence, ($29_1$) becomes

$$\sum_{m=1}^{n} d(m) = 2n\sum_{m=1}^{n^{\frac{1}{2}}} 1/m - [n^{\frac{1}{2}}]^2 + O(n^{\frac{1}{2}}) = 2n(\log n^{\frac{1}{2}} + C) - n + O(n^{\frac{1}{2}}),$$

which is Dirichlet's estimate in his divisor problem.

Next, suppose that $F'(n) = O(1)$. Then it is clear from (28) that the assumption, $f'(n) = O(1)$, of ($29_1$) is satisfied. But ($29_1$) now implies that

$$\sum_{m=1}^{n} f(m) = n\sum_{m=1}^{n^{\frac{1}{2}}} f'(m)/m + O(n^{\frac{1}{2}}) \equiv n\sum_{m=1}^{\infty} f'(m)/m + O(n^{\frac{1}{2}}). \tag{$29_2$}$$

In fact, it is seen from (28) by partial summation that, since $F'(n) = O(1)$, the infinite series $\Sigma f'(m)/m$ is convergent and is approximated by its $n$-th partial sum with an error $O(n^{-1})$.

For instance, if $f'(n)$ is $(-1)^{\frac{1}{2}(n-1)}.4$ or 0 according as $n$ is odd or even, then $f(n)$ is, by (15), the number, $r_2(n)$, of the representations of $n$ as a sum of two squares (cf. §71 below). Hence, ($29_2$) is reduced to

$$\sum_{m=1}^{n} r_2(m) = 4n\sum_{m=1}^{\infty} (-1)^{m-1}/(2m-1) + O(n^{\frac{1}{2}}) \equiv \pi n + O(n^{\frac{1}{2}}),$$

which is Gauss' estimate in his lattice problem.

The assumption of ($29_2$) was the boundedness of the partial sums, (28), of the series $\Sigma f'(m)$. Suppose now that this series is convergent. Then it is clear from the passage from ($29_1$) to ($29_2$), that ($29_2$) can be refined to

$$\sum_{m=1}^{n} f(m) = n\sum_{m=1}^{n} f'(m)/m + o(n^{\frac{1}{2}}) \equiv n\sum_{m=1}^{\infty} f'(m)/m + o(n^{\frac{1}{2}}). \tag{$29_3$}$$

For instance, if $f(n) = |\mu(n)|$, the convergence of $\Sigma f'(n)$ is readily seen to mean the convergence of $\Sigma\, \mu(n)/n$ (cf. the first part of §53 below), and so $(29_3)$ represents that restatement of the prime number theorem to which Landau devoted pp. 604–609 of his *Handbuch*.

The assumption of $(29_3)$ neither implies, nor is implied by, the estimate $f'(n) = O(1/n)$; an estimate which does not even entail the estimate $F'(n) = O(1)$, supposed by $(29_2)$. However, if $f'(n) = O(1/n)$, it is clear from (20) that

$$(29_4) \qquad \sum_{m=1}^{n} f(m) = n\sum_{m=1}^{n} f'(m)/m + \sum_{m=1}^{n} O(1/m) \equiv n\sum_{m=1}^{\infty} f'(m)/m + O(\log n).$$

Since $f'(n) = O(1/n)$, it is seen from $(29_4)$ by partial summation that

$$(29_4 \text{ bis}) \qquad \sum_{m=1}^{n} mf(m) = \tfrac{1}{2}n^2 \sum_{m=1}^{\infty} f'(m)/m + nO(\log n).$$

On the other hand, since $f'(n) = O(1/n)$, partial summation of $(29_4$ bis) gives only $O(\log^2 n)$, instead of $O(\log n)$, in $(29_4)$. Thus the true estimate belonging to the assumption $f'(n) = O(1/n)$ is $(29_4)$, and not $(29_4$ bis).

For instance, if $f'(n) = 1/n$ or $f'(n) = \mu(n)/n$, then (15) shows that $f(n) = \sigma(n)/n$ and $f(n) = \phi(n)/n$ respectively, where $\sigma(n)$ denotes the sum of the divisors of $n$ and $\phi(n)$ is Euler's function. Thus $(29_4$ bis) is reduced to

$$\sum_{m=1}^{n} \sigma(m) = \tfrac{1}{2}n^2\zeta(2) + nO(\log n) \quad \text{and} \quad \sum_{m=1}^{n} \phi(m) = \tfrac{1}{2}n^2/\zeta(2) + nO(\log n),$$

the standard estimates of Dirichlet and Mertens. In contrast to the preceding estimates of classical sum functions, neither of the latter estimates has ever been improved. On the other hand, it has never been proved in either case that the $O(\log n)$ is not $o(\log n)$.

## The Pair of Eratosthenian Summation Methods

**8.** Let $f(n)$ be arbitrary. Since $[n/m]/n \to 1/m$ as $n \to \infty$ if $m$ is fixed, (19) implies the formal connection

$$(30) \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} f(m) \equiv \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} \left[\frac{n}{m}\right] f'(m) = \sum_{m=1}^{\infty} \frac{1}{m} f'(m),$$

if the limit process is carried out term-by-term.

The evaluation of $M(f)$ resulting from (3), (30) in the series form $\Sigma f'(m)/m$ appears to have been known to Gauss, who has used it freely (for instance, in his celebrated fragment [27] which, having been posthumous, was not at the disposal of the parallel and more inclusive work of Dirichlet; in fact, Dirichlet [16] knew only of what the *Disquisitiones Arithmeticae* [26] stated without proof concerning certain averages).

Correspondingly, the sieve process that Dedekind [14], in his commentaries to Gauss' fragment, calls the successive elimination of primes, is nothing but (15), followed by the formal limit process applied in (30). Dedekind also

observes that the elimination of *all* the primes, that is, the limit process corresponding to the formal step

$$\lim_{n\to\infty} \sum_{m=1}^{n} \frac{1}{n}\left[\frac{n}{m}\right] f'(m) = \sum_{m=1}^{\infty} \lim_{n\to\infty} \frac{1}{n}\left[\frac{n}{m}\right] f'(m) \tag{31}$$

in (30), needs legitimization. In the particular cases considered by Gauss, cases in which $f'(n)$ is substantially a non-principal character, Dedekind justifies this step by a partial summation (actually, the assumption of $(29_2)$ is satisfied in these cases).

Cesàro has rediscovered the Gaussian principle and, being free of Dedekindian inhibitions as to (31), has applied (30) extensively; cf. [9], [10].

According to (3), the claim made by (30) is that $M(f) = \Sigma f'(m)/m$ holds whenever either side of the latter equation exists. This claim is three-fold:

(I) $M(f) \neq \Sigma f'(m)/m$ is impossible if the mean $M(f)$ exists and the series $\Sigma f'(m)/m$ is convergent;

(II) $M(f)$ must exist if $\Sigma f'(m)/m$ converges;

(III) $\Sigma f'(m)/m$ must converge if $M(f)$ exists.

The object of this chapter is to show that, though (I) happens to be true, both (II) and (III) are false. It will be seen in §13 and §16 that even the truth of (I) is due only to a fortunate coincidence.

**9.** If the $f(n)$ and $f'(n)$ are replaced by the functions

$$F(n) = \frac{1}{n}\sum_{m=1}^{n} f(m) \quad \text{and} \quad F^*(n) = \sum_{m=1}^{n} \frac{f'(m)}{m}, \tag{32}$$

then the relation (15) defining the connection between $f(n)$ and $f'(n)$ is equivalent to

$$F(n) = \sum_{m=1}^{n} \left\{\left[\frac{n}{m}\right]\frac{m}{n} - \left[\frac{n}{m+1}\right]\frac{m+1}{n}\right\} F^*(m). \tag{33}$$

In fact, substitution of (32) into (19) gives

$$F(n) = \frac{1}{n}\sum_{m=1}^{n}\left[\frac{n}{m}\right] m\{F^*(m) - F^*(m-1)\},$$

a relation identical with (33), since $\left[\dfrac{n}{n+1}\right] = 0$ and $F^*(0) = 0$, by (32).

The following considerations depend on the fact that (33) is a linear transformation,

$$F(n) = \sum_{m=1}^{n} a_{nm} F^*(m) \quad \text{or} \quad F(n) = \sum_{m=1}^{\infty} a_{nm} F^*(m), \tag{34}$$

belonging to the matrix formed by the absolute constants

$$a_{nm} = \left[\frac{n}{m}\right]\frac{m}{n} - \left[\frac{n}{m+1}\right]\frac{m+1}{n}, \quad (a_{nm} = 0 \text{ for } m > n). \tag{35}$$

According to (32), the existence of $M(f)$ and the convergence of $\Sigma f'(n)/n$ are equivalent to the existence of the limits, $F(\infty)$ and $F^*(\infty)$, respectively. It follows therefore from (34) that the investigation of any connection between the existence of $M(f)$ and the convergence of $\Sigma f'(n)/n$ must be reducible, at least in principle, to the general theory of linear summation methods.

The simplest result attainable by this approach is as follows:

(I) *If $M(f)$ exists and if $\Sigma f'(n)/n$ is convergent, then* $M(f) = \Sigma f'(n)/n$.

In view of the remark following (35), the assertion of (I) is that the summation process (34) defined by the matrix (35) satisfies the conditions of consistency. The latter are known to be two-fold, namely

$$\text{(i)} \qquad \lim_{n\to\infty} \sum_{m=1}^{n} a_{nm} = 1; \qquad\qquad \text{(ii)} \qquad \lim_{n\to\infty} a_{nm} = 0 \text{ for } m = 1, 2, \cdots .$$

But it is clear from (35) that the sum occurring in (i) is 1 for every $n$, and that (ii) also is satisfied, since $[x]/x \to 1$ as $x \to \infty$.

**10.** On p. 115 of his book, Cesàro [9] attempted a short proof of a general theorem which, in terms of the notations (3) and (15), can be formulated as follows: If the series $\Sigma f'(n)/n$ is convergent and if $f'(1) + \cdots + f'(n) = o(n)$, then $M(f)$ exists. However, his proof of the theorem is readily seen to be erroneous. It turns out that the theorem itself is false.

First, the assumptions of the theorem are tautological, since, according to (4), the $o$-condition of Cesàro is always implied by the convergence of $\Sigma f'(n)/n$. Hence, what the theorem actually states is that the convergence of $\Sigma f'(n)/n$ is sufficient for the existence of $M(f)$. In view of (I) and of the remark following (35), this can be expressed by saying that the summation process (34) defined by the matrix (35) is regular in the sense of the theory of Toeplitz ([78]; cf. Schur [70]). But then the sums corresponding to Lebesgue's constants in the theory of Fourier series must be bounded. In other words, not only the conditions (i), (ii) of (I) but also the third condition of Toeplitz, namely

$$\text{(iii)} \qquad \sum_{m=1}^{n} |a_{nm}| = O(1) \quad \text{as} \quad n \to \infty,$$

must be satisfied by (35). Conversely, it is clear from (I) that (iii) is not only necessary but sufficient as well for the truth of Cesàro's theorem.

It will now be shown that (iii) is violated by (35).

(II) *The convergence of $\Sigma f'(n)/n$ does not imply the existence of $M(f)$.*

Actually, it turns out that the sum (iii) belonging to (35) exceeds a constant multiple of $\log n$. Incidentally, it will be seen from the proof that (iii), just as in the theory of Fourier series, becomes true if $O(1)$ is replaced by $O(\log n)$.

Let $n$ have a fixed value. It is clear that, if $k$ is a positive integer not greater than $n$, a positive integer $m$ satisfies the pair of conditions

$$\left[\frac{n}{m}\right] = k, \qquad \left[\frac{n}{m+1}\right] = k$$

if and only if it fulfills both inequalities

$$\frac{n}{k+1} < m \leqq \frac{n}{k} - 1.$$

Hence, if these inequalities are satisfied, then (35) can simply be written as

$$a_{nm} = k\frac{m}{n} - k\frac{m+1}{n},$$

which implies that $|a_{nk}| = k/n$. Accordingly, if $l_{nk}$ denotes the number of the integers $m$ satisfying both inequalities which belong to a given $k$ and to the fixed $n$, then, since $k = 1, 2, \cdots, n$,

$$\sum_{m=1}^{n} |a_{nm}| \geqq \sum_{k=1}^{n} l_{nk}k/n \geqq \sum_{k=1}^{n^{\frac{1}{2}}} l_{nk}k/n.$$

It is understood that the first of the two signs $\geqq$ is introduced by the omission of those summation indices which do not satisfy any of the pair of inequalities belonging to the various values of $k$ and to the fixed $n$ (and that the summation limit $n^{\frac{1}{2}}$ refers to $k = [n^{\frac{1}{2}}]$). But it is clear from the definition of $l_{nk}$ that

$$l_{nk} \geqq \frac{n}{k} - 1 - \frac{n}{k+1} - 1 = \frac{n}{k(k+1)} - 2,$$

which implies that $l_{nk}k/n \geqq (k+1)^{-1} - 2k/n$. Hence

$$\sum_{m=1}^{n} |a_{nm}| \geqq \sum_{k=1}^{n^{\frac{1}{2}}} \{(k+1)^{-1} - 2k/n\} > \sum_{k=2}^{n^{\frac{1}{2}}} k^{-1} - 2\sum_{k=1}^{n^{\frac{1}{2}}} k/n,$$

and so, since $2\sum_{k=1}^{x} k = O(x^2)$ as $x \to \infty$,

$$\sum_{m=1}^{n} |a_{nm}| > \sum_{k=2}^{n^{\frac{1}{2}}} k^{-1} - O(n)/n = \log(n^{\frac{1}{2}}) + O(1)$$

as $n \to \infty$. Since this contradicts (iii), the proof of (II) is complete.

**11.** Since (35) implies that $a_{nn} = 1$, the transformation (34) of $F^*$ into $F$ has a unique inverse, say

$$F^*(n) = \sum_{m=1}^{n} b_{nm}F(m) \quad \text{or} \quad F^*(n) = \sum_{m=1}^{\infty} b_{nm}F(m), \tag{36}$$

where $b_{nm} = 0$ for $m > n$. Corresponding to (35), the infinite matrix $(b_{nm}) = (a_{nm})^{-1}$ consists of certain absolute constants. Some of the $b_{nm}$ are negative, as are, by (35), some of the $a_{nm}$; so that neither of the summation methods (34), (36) belongs to "weights".

It will now be shown that what (II) states for (34) holds for (36) also.

(III) *The existence of $M(f)$ does not imply the convergence of $\Sigma f'(n)/n$.*

In order to prove (III), it would be possible to proceed in the same way as in

§10, that is, to show that the Lebesgue-Toeplitz condition (iii), §10 remains violated if $a_{nm}$ is replaced by $b_{nm}$. However, it is seen from (36), (32) and (24) that the explicit representation of $b_{nm}$ involves awkward combinations of the Möbius function. It is therefore convenient to carry out in another form the considerations on which the Lebesgue-Toeplitz criterion depends. To this end, it is sufficient to observe that, according to (4), the negation (III) is a corollary of the first part of the following statement:

(III*) *The convergence of* $\Sigma f(n)/n$ *is insufficient and unnecessary for the convergence of* $\Sigma f'(n)/n$.

The truth of the second part of (III*), which is not needed for (III), follows by choosing $f'(n) = 0$ for every $n > 1$, since $f(n)$ then is the constant $f'(1)$, by (15). In order to prove the first part of (III*), it is sufficient to observe that, according to (25), the series $\Sigma f'(n)/n$ is the Dirichlet product of the series $\Sigma \mu(n)/n$ and $\Sigma f(n)/n$ (in fact, (25) and the definition (22) imply that the Dirichlet series of $1/\zeta(s)$ is $\Sigma \mu(n)n^{-s}$). Thus the assertion of the first part of (III*) is that there exist convergent numerical series whose Dirichlet product with the series $\Sigma \mu(n)/n$ is a divergent series. But $\Sigma \mu(n)/n$ is not absolutely convergent, since $\Sigma \mid \mu(n) \mid /n \geqq \Sigma \mid \mu(p) \mid /p$, where $\mu(p) = -1$ for every prime $p$. Hence, the existence of numerical series of the type required follows from the converse of the multiplication theorem of Stieltjes; a converse established by Schur [70] exactly along the lines of the Lebesgue-Toeplitz principle.

**12.** It is seen from the remarks following (35), that (III), (II) and (I) together can be expressed as follows: The linear summation methods defined by the *reciprocal* transformations (34), (36) belonging to (35) are *incomparable*, though *consistent.*

The standard instances of consistent but incomparable summation methods are supplied by the processes of Cesàro-Hölder and of Euler or of Borel. However, the present instance of incomparable consistency is more striking, since the processes (34) and (36), being reciprocal, form an *involutory* pair. However, as pointed out at the beginning of §11, the "weights" are now negative in part.

It seems to be of particular interest that the pathological situation represented by the incomparable character of the consistent involution (34), (36) is inherent in the sieve of Eratosthenes (cf. (30), (31) and §5). In fact, it is clear from (32) that the reciprocal mates (34), (36) represent a mere transcription of the classical pair (15), (24). Thus the "pathology" cannot now be excused by the "artificial" nature of the analytic smoothing processes of Euler or of Borel.

## Euler Products

**13.** Most of the functions $f(n)$ considered in the classical literature (cf. §7–§8) are multiplicative, that is to say such that, if $(n_1, n_2)$ denotes the greatest common divisor of $n_1$ and $n_2$,

$$f(n_1 n_2) = f(n_1)f(n_2) \quad \text{if} \quad (n_1, n_2) = 1. \tag{37}$$

This implies that, if the trivial case $f(1) = f(2) = \cdots = 0$ is excluded,

(38) $$f(1) = 1; \quad \text{hence} \quad f'(1) = 1, \text{by (18)}.$$

According to (37), a multiplicative $f(n)$ is uniquely determined for every $n$ by an arbitrary assignment of a double sequence of values $f(p^k)$, where $p$ and $k$ range over all primes and over all positive integers respectively. It is easily verified from (15) that

(39) $$f \text{ is multiplicative if and only if } f' \text{ is.}$$

Thus (15) or (24) is now equivalent to the identity

(40) $$f(p^k) = \sum_{j=0}^{k} f'(p^j), \quad \text{i.e.} \quad f'(p^k) = f(p^k) - f(p^{k-1}), \qquad k > 0,$$

(which, according to (17), is true, but incomplete, if (37) is omitted).

If $f(n)$ is multiplicative, the same is true, by (39), of $f'(n)/n$. Hence, if $\Sigma f'(n)/n$ is convergent, a formal application of Euler's factorization gives

(41) $$\sum_{n=1}^{\infty} f'(n)/n = \prod_p \left\{1 + \sum_{k=1}^{\infty} f'(p^k)/p^k\right\};$$

cf. (50) below. But then (40) and (38) show that the Gaussian principle $M(f) = \Sigma f'(n)/n$ (cf. §8) can be written in the form

(42) $$M(f) = \prod_p \sum_{k=0}^{\infty} \{f(p^k) - f(p^{k-1})\}/p^k, \qquad f(1) = 1, \quad \text{if} \quad f(p^{-1}) = 0.$$

(42) admits of an interpretation in terms of the product rule of "independent probabilities" (cf. Wintner [92]). In fact, let $f_q(n)$ denote, for a fixed prime $q$, the multiplicative function for which the double sequence $\{\{f_q(p^k)\}\}$ is given by

(43) $$f_q(p^k) = 1 \quad \text{if} \quad p \neq q, \quad \text{and} \quad f_q(q^k) = f(q^k),$$

where $k = 1, 2, \cdots$. Then, if (42) is applied to $f = f_q$, it follows that (42) itself can be written in the form

(44) $$M(f) = \prod_p M(f_p).$$

On the other hand, from (43), (38) and (37),

(45) $$f(n) = \prod_p f_p(n) \quad \text{for every } n$$

(it being understood that the infinite product (45) has only a finite number of factors distinct from 1, if $n$ is fixed). It is the formal parallelism between (44) and (45) on which certain considerations of Sylvester [73], [76] appear to depend.

Actually, (42) presupposes not only the validity of the Gaussian principle (§8) in the multiplicative case, a principle disproved by (II) and (III) in the general case, but also the legitimacy of the Euler factorization (41). It will now

be shown that (42) can fail in *every* sense, that is, that *all three* statements representing the Eulerian analogues of (I), (II), (III), §8 are false.

**14.** A realization of the first of these three contingencies can be made to depend on the machinery of the prime number theorem, as follows:

The proto-Tauberian direct proof of $\mu(1) + \cdots + \mu(n) = o(n)$, where $\Sigma\, \mu(n)n^{-s} = 1/\zeta(s)$, is based on an estimate of the function $1/\zeta(s)$ on a domain containing the line $\sigma = 1$. Since the same estimate is trivial for the function $\zeta(s)$ itself and therefore for the function $\zeta(2s - 1)$ as well, it is available for the product of $1/\zeta(s)$ and $\zeta(2s - 1)$ also. Hence, if $1/\zeta(s)$ is replaced by $\zeta(2s - 1)/\zeta(s)$ in the proof of $\mu(1) + \cdots + \mu(n) = o(n)$, it follows that the sum of the first $n$ coefficients in the Dirichlet series of $\zeta(2s - 1)/\zeta(s)$ is $o(n)$. This implies that the latter Dirichlet series is convergent at $s = 1$, since the function $\zeta(2s - 1)/\zeta(s)$ is regular at $s = 1$ (Fatou-M. Riesz; cf. §24 below).

It is convenient to restate this result in a simpler form, as follows: If $g(n)$ is defined by

$$(46)\qquad \sum g(n)n^{-s} = G(s), \quad \text{where} \quad G(s) = \prod_p \left\{1 - \frac{1}{p^s} + \frac{p}{p^{2s}}\right\}, \qquad (\sigma > 1),$$

then

$$(47)\qquad g(1) + \cdots + g(n) = o(n).$$

In fact, two-fold application of $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ gives

$$(48)\qquad \frac{\zeta(2s-1)}{\zeta(s)} = \prod_p \left\{1 - \frac{1}{p^s} + \frac{p}{p^{2s}} - \frac{p}{p^{3s}} + \frac{p^2}{p^{4s}} - \frac{p^2}{p^{5s}} + \cdots\right\},$$

where $\sigma > 1$. Clearly, the quotient of the products (46), (48) is absolutely convergent, and represents a non-vanishing regular function, in a half-plane containing the line $\sigma = 1$ in its interior. In particular, the quotient has a Dirichlet series that is absolutely convergent at $s = 1$. Since the Dirichlet series of the function (48) is convergent at $s = 1$, it follows from the multiplication theorem of Stieltjes, that the Dirichlet series of the function (46) is convergent at $s = 1$. In other words, $\Sigma\, g(n)/n$ is convergent. Hence, (47) follows from (4).

It should be noted for later reference that the function $\zeta(2s - 1)/\zeta(s)$ attains at $s = 1$ the value

$$(49)\qquad \{\zeta(2s - 1)/\zeta(s)\}_{s=1} = \tfrac{1}{2},$$

since $(z - 1)\zeta(z) \to 1$ as $z \to 1$ and $(s - 1)/(2s - 1 - 1) = \frac{1}{2}$.

**15.** According to Euler [24], the formal factorization belonging to an arbitrary multiplicative function $g(n)$ is

$$(50)\qquad \sum g(n) = \prod_p \left\{1 + \sum_{k=1}^{\infty} g(p^k)\right\}.$$

If $g \geqq 0$, then (50) is valid whether $\Sigma\, g(n) < \infty$ or $\Sigma\, g(n) = \infty$. Hence

$$\sum |g(n)| < \infty \quad \text{if and only if} \quad \sum_p \sum_{k=1}^{\infty} |g(p^k)| < \infty. \tag{51}$$

If $\Sigma\, g(n)$ is absolutely convergent, then (50) is valid and, in addition, the product and each of the sums on the right of (50) are absolutely convergent. However, all of this can hold even if $\Sigma\, g(n)$ is divergent. This is shown by the example

$$g(p) = -1, \qquad g(p^2) = 1, \qquad g(p^k) = 0 \quad \text{if} \quad k > 2. \tag{52}$$

In fact, every factor $\{\ \}$ in (50) is then 1, although not even $g(n) \to 0$ is satisfied. On the other hand, if

$$g(p^k) = (-1)^{p-1} p^{-\frac{1}{2}k}, \tag{53}$$

then $\Sigma\, g(n)$ is convergent, since $g(n) = (-1)^{n-1} n^{-\frac{1}{2}}$, although the product (50) becomes

$$\left\{1 - \sum_{k=1}^{\infty} 2^{-\frac{1}{2}k}\right\} \prod_{p>2} \left\{1 + \sum_{k=1}^{\infty} p^{-\frac{1}{2}k}\right\} = -\infty.$$

The examples (52) and (53) prove the second and the third of the three negations italicized at the end of §13; negations which can be formulated as follows:

(IV) *If $g(n)$ is a multiplicative function, then*

(I′) *the convergence of $\Sigma\, g(n)$ and the convergence of its Euler product do not imply the equality* (50);

(II′) *the convergence of $\Sigma\, g(n)$ does not imply the convergence of its Euler product*;

(III′) *the convergence of $\Sigma\, g(n)$ is not implied by the convergence of its Euler product*;

*and all of this holds even if each of the series $\{\ \}$ on the right of* (50) *is required to converge absolutely.*

(I′) can be deduced from §14 by choosing

$$g(p^{2k}) = p^{-k}, \qquad g(p^{2k-1}) = -p^{-k} \tag{54}$$

(cf. also §55 and §64 below). In fact, it is clear from (54) that the Euler factorization of the Dirichlet series $\Sigma\, f(n)n^{-s}$ belonging to $f(n) = ng(n)$ is identical with the product (48), where $\sigma > 1$. Hence the result of §14 means that $\Sigma\, g(n)$ is a convergent series. According to (49), the value represented by $\Sigma\, g(n)$ is $\frac{1}{2}$, since the function $\Sigma\, ng(n)n^{-s} = \zeta(2s-1)/\zeta(s)$ is continuous at $s = 1$ and so, by Dirichlet's analogue of Abel's continuity theorem,

$$\Sigma\, g(n)n^{-\epsilon} \to \Sigma\, g(n) \quad \text{as} \quad \epsilon \to +0. \tag{55}$$

On the other hand, (54) shows that the factor $\{\ \}$ of the product (50) is $1 - p^{-1} + p^{-1} - p^{-2} + p^{-2} - \cdots = 1$ for every $p$. Hence (50) claims that $\frac{1}{2} = 1$.

**16.** It is clear from this proof that what is responsible for the first item of (IV) is the failure of the analogue of the Abelian theorem (55) in the case in which, under the assumption that $g(n)$ is multiplicative, both series (55) are replaced by their Euler products. In fact, what is accomplished by (54) is that the Euler product

$$P(\epsilon) = \prod_p \left\{1 + \sum_{k=1}^{\infty} g(p^k)p^{-k\epsilon}\right\} \tag{56}$$

of the Dirichlet series $\Sigma\, g(n)n^{-\epsilon}$ converges both for $\epsilon > 0$ and for $\epsilon = 0$, and the limit $P(+0)$ exists, but $P(+0) \neq P(0)$. This situation was met by Hardy [35] in a case numerically more involved than, but otherwise of exactly the same type as, the example (54); cf. the end of §55 below.

A possibility distinct from the one just described could be this: The Euler product (56) of $\Sigma\, g(n)n^{-\epsilon}$ converges for both $\epsilon > 0$ and $\epsilon = 0$ but the analogue of the Abelian theorem fails for the reason that the limit $P(+0)$ does not exist.

Actually, the discrepancy between the behavior of $\Sigma\, g(n)n^{-s}$ and of its Euler factorization $P(s)$ is so far-reaching that $P(s)$ need not even have an abscissa of convergence. In other words, if $P(s)$ converges at $s = 1$, it need not converge when $\sigma > 1$. In fact, it is possible to have for $P(\epsilon)$ an isolated convergence point. For instance, if

$$g(p) = p, \qquad g(p^2) = -p^2, \qquad g(p^k) = 0 \quad \text{if} \quad k > 2, \tag{57}$$

then every factor $\{\ \}$ of (56) is 1 at $\epsilon = 1$, although (56) is divergent when either $1 < \epsilon \leqq 2$ or $0 \leqq \epsilon < 1$.

It would be interesting to know whether an Euler product $P(s)$ can or cannot overconverge its Dirichlet series; in the sense that the product $P(s)$ converges, and represents a regular function, in a half-plane $\sigma > \theta$, where $\theta$ is less than the convergence abscissa of $\Sigma\, g(n)n^{-s}$.

**17.** The above counter-examples depend on a retroaction of what is contributed to (50) by those prime powers, $p^2, p^3, \cdots$, which are not primes, $p$; a retroaction excluded in standard situations, characterized by

$$\prod_p \sum_{k=2}^{\infty} |\, g(p^k)\,| < \infty. \tag{58}$$

If (58) is satisfied, the classical condition, (51), for the validity of (50) requires the absolute convergence of $\Sigma\, g(p)$. This requirement is rather crude, since it excludes, among other things, every case in which (50) is valid but either the product (50) or the series $\Sigma\, g(n)$ is conditionally convergent. A criterion free of this objection may be formulated as follows:

(V) *If $g(n)$ is a multiplicative function satisfying $\Sigma\, |\, g(p)\,|^2 < \infty$ and $g(p^2) = g(p^3) = \cdots 0$, then either the product* (50) *and the series $\Sigma\, g(n)$ are convergent and have the same value or both the product* (50) *and the series $\Sigma\, g(n)$ are divergent, according as the series $\Sigma\, g(p)$ is convergent or divergent.*

The assumption $g(p^k) = 0$ for $k > 1$ is made only for the sake of simplicity; it could be replaced by a summary assumption corresponding to (58). As it stands, it simplifies the product (50) to $\Pi(1 + a_i)$, if $a_i = g(p)$, where $p = p_i$ denotes the $i$-th prime. It follows therefore from the other assumption, $\Sigma \mid g(p) \mid^2 < \infty$, of (V) and from Cauchy's criterion for the convergence of a product $\Pi(1 + a_i)$, that the convergence of the product (50) is equivalent to the convergence of the series $\Sigma\, g(p)$.

Thus it is clear that, in order to prove (V), it is sufficient to verify that $\Sigma \mid g(p) \mid^2 < \infty$ implies $\Sigma' \mid g(n) \mid < \infty$, where the accent refers to the omission of those summation indices $n$ which are primes. But since $g(p^k) = 0$ for $k > 1$, the indices $n$ occurring in $\Sigma' g(n)$ are those square-free integers which are not primes. Accordingly, if $(p_1, \cdots, p_j)$ runs through all $j$-tuples of distinct primes in the sums

$$s_j = \Sigma \mid g(p_1) \cdots g(p_j) \mid \quad \text{and} \quad t_j = \Sigma \mid g(p_1) \cdots g(p_j) \mid^2,$$

then the assumption and the assertion are $t_1 < \infty$ and $s_2 + s_3 + s_4 + \cdots < \infty$ respectively. But $s_2 + s_3 + s_4 + \cdots < \infty$ is readily seen to be implied by $t_1 < \infty$.

In fact, the Schwarz inequality shows that $s_j \leqq (t_1 t_{j-1})^{\frac{1}{2}}$ for every $j > 1$, and it is obvious that $t_j \leqq t_1 t_{j-1}$. Hence $s_2 + s_3 + s_4 + \cdots$ is majorized by a convergent geometric series, if $t_1 < 1$. But the assumption $t_1 < 1$ involves no loss of generality. This is seen by observing that $t_1 = \Sigma \mid g(p) \mid^2$, and that the assumption $\Sigma \mid g(p) \mid^2 < \infty$ of (V) can be replaced by $\Sigma \mid h(p) \mid^2 < \infty$, if $h(n)$ is a function satisfying $h(p) = g(p)$ for all but a finite number of primes, and $h(p) = 0$ for the remaining primes. Since the replacement of $g(n)$ by such an $h(n)$ cannot affect the truth of (V), the proof of (V) is complete.

## Tauberian Theorems for the Eratosthenian Summation Methods

**18.** For an arbitrary $f(n)$ or $f'(n)$, the functions $F(n)$, $F^*(n)$ defined by (32) represent the $n$-th approximations to the limits $F(\infty) = M(f)$, $F^*(\infty) = \Sigma f'(n)/n$, if any. According to (34) and (36), the transition from either of the functions $F(n)$, $F^*(n)$ to the other depends on a linear summation process. And (I) states that these two Eratosthenian summation processes are consistent. On the other hand, (II) and (III) state that Eratosthenian summability of either kind does not imply Eratosthenian summability of the other kind. Thus there arises the need for Tauberian restrictions under which Eratosthenian summability of the first kind implies Eratosthenian summability of the second kind, and for Tauberian restrictions under which the converse is true. Such Tauberian theorems form the subject of this chapter.

It turns out that Tauberian theorems of the one kind (those leading from $\Sigma f'(n)/n = \lim F^*$ to $M(f) = \lim F$) are, in the main, elementary in nature, while the principal Tauberian theorem of the other kind (leading from $M(f) = \lim F$ to $\Sigma f'(n)/n = \lim F^*$) involves the prime number theorem. Actually, the latter Tauberian theorem implies an extension of the prime number theorem to the case of an arbitrary sequence of primes; cf. §77 below.

**19.** It is easily verified that, corresponding to the formal identity (25) between Dirichlet series, the relation (15) defining the connection between $f$ and $f'$ is formally equivalent to the Lambertian identity

$$\Sigma f(n)r^n = \Sigma f'(n)r^n/(1 - r^n), \tag{59}$$

where the range of the summation indices is that specified by (1). The formal derangements identifying the two series (59) also show that either both the power series (59) and the Lambert series (59) are, or neither of these series is, convergent (or, what is the same thing, absolutely convergent) for $r < 1$.

Due to (59), both Tauberian problems of §18 become reducible to classical Tauberian problems in the summation theory of series.

First, it is clear from (3) that $M(f)$ exists if and only if the series

$$\Sigma\, h(n), \quad \text{where} \quad h(n) = f(n) - f(n-1) \quad \text{if} \quad n > 1 \quad \text{and} \quad h(1) = f(1), \tag{60}$$

is summable $(C, 1)$. Furthermore, since the $n$-th partial sum of the series (60) is $f(n)$,

$$(1 - r)^{-1} \sum h(n)r^n = \sum f(n)r^n. \tag{61}$$

Hence, the Abelian summability $(A)$ of the series (60) means that the power series $\Sigma f(n)r^n$ converges for $r < 1$ and that the product

$$(1 - r)\Sigma f(n)r^n \tag{62}$$

tends to a limit as $r \to 1$. Correspondingly, since $(1 - r)r^n/(1 - r^n) \to 1/n$ as $r \to 1$, the Lambertian summability $(L)$ of a series, say of $\Sigma\, g(n)$, is defined by the property that the Lambert series $\Sigma\, ng(n)r^n/(1 - r^n)$ converges for $r < 1$ and that the product

$$(1 - r) \sum ng(n)r^n/(1 - r^n) \tag{63}$$

tends to a limit as $r \to 1$.

The fundamental theorem on the three summation processes $(C, 1)$, $(L)$, $(A)$ is that summability $(L)$ implies summability $(A)$ and is implied by summability $(C, 1)$ [and, incidentally, even by summability $(C, N)$]. The first of these two facts, which is due to Hardy and Littlewood [37], lies somewhat deeper than the prime number theorem. The second statement is elementary; cf. Hardy [34]. Any two of the three summation processes are consistent and inequivalent.

**20.** A few elementary facts will now be collected.

(VI) *If* $|f'(1)| + \cdots + |f'(n)| = O(n)$, *and if* $M(f')$ *exists, then*

$$\frac{1}{n}\sum_{m=1}^{n} f(m) - \sum_{m=1}^{n}\frac{f'(m)}{m} \to -\Gamma'(2)M(f') \quad \text{as} \quad n \to \infty. \tag{64}$$

According to (2), the $O$-assumption of (VI) can be written in the form $N(f') < \infty$. Hence, the assumptions of (VI) are precisely those made for $g = f'$ before (9). But (19) shows that (64) is identical with the case $g = f'$ of (9).

In view of the case $g = f'$ of (4), the Axerian lemma (VI) contains the following corollary, which is Tauberian with reference to (II).

(VII) *If* $|f'(1)| + \cdots + |f'(n)| = O(n)$, *then the convergence of* $\Sigma f'(n)/n$ *implies the existence of* $M(f)$.

(VII) is a Tauberian refinement of one part of the following criterion, both parts of which are trivial from (20) and (3).

(VIII) *If* $|f'(1)| + \cdots + |f'(n)| = o(n)$, *then the convergence of* $\Sigma f'(n)/n$ *is equivalent to the existence of* $M(f)$.

It is clear from (VIII) and from the case $g = |f'|$ of (4), that the absolute convergence of $\Sigma f'(n)/n$ is sufficient for the existence of $M(f)$. However, it will be seen in §27 and §33 that the existence of $M(f)$ is not the true consequence of the absolute convergence of $\Sigma f'(n)/n$.

**21.** The proof of the following criterion, a criterion Tauberian with reference to (III), will involve the prime number theorem (cf. §19).

($IX_1$) *If* $f'(n) = O_L(1)$, *then the existence of* $M(f)$ *implies the convergence of* $\Sigma f'(n)/n$.

In fact, suppose that $M(f)$ exists. This means that the series (60) is summable $(C, 1)$, and so it is summable $(A)$; so that the product (62) tends to a limit as $r \to 1$. It follows therefore from (59) that the product (63) belonging to $g(n) = f'(n)/n$ tends to a limit as $r \to 1$. In other words, the series $\Sigma f'(n)/n$ is summable $(L)$. Consequently, it is summable $(A)$. Hence, in order to complete the proof of ($IX_1$), it is sufficient to apply to the series $\Sigma f'(n)/n$ an elementary Tauberian theorem of Hardy and Littlewood, according to which a series is convergent whenever it is summable $(A)$ and such that its $n$-th term is $O_L(1/n)$; cf. Karamata [52].

The deep, though Abelian (that is, not Tauberian) theorem according to which summability $(L)$ always suffices for summability $(A)$ will not be used in the proof of the following fact, which represents a *dual*, instead of being a *converse*, of ($IX_1$), and is Tauberian with reference to (II).

($IX_2$) *If* $f(n) = O_L(1)$, *then the convergence of* $\Sigma f'(n)/n$ *implies the existence of* $M(f)$.

In fact, suppose that the series $\Sigma f'(n)/n$ is convergent. Then it is summable $(L)$. This means that the product (63) belonging to $g(n) = f'(n)/n$ tends to a limit as $r \to 1$. It follows therefore from (59) that the product (62) tends to a limit as $r \to 1$. In other words, the series (60) is summable $(A)$. On the other hand, the existence of $M(f)$ means that the series (60) is summable $(C, 1)$. Hence, in order to complete the proof of ($IX_2$), it is sufficient to apply to the series (60) an elementary Tauberian theorem of Hardy and Littlewood, according to which a series is summable $(C, 1)$ whenever it is summable $(A)$ and such that its $n$-th partial sum is $O_L(1)$; cf. Karamata [52].

($IX_2$) applies in the same direction as (VII) but is independent of (VII). In fact, even if $O_L(1)$ is replaced by $O(1)$ in ($IX_2$), all that results from (24) is $f'(n) = O(d(n))$, where $d(n)$ denotes the number of the divisors of $n$. Since $d(1)$

$+ \cdots + d(n) \sim n \log n$, it follows that the $O$-condition of (VII) is short by the factor $\log n$ (which, as a matter of fact, is the *same* $\log n$ as the one mentioned after the italicized result in §10).

Incidentally, it is clear from the proof of ($IX_2$) that the assertion of ($IX_2$) remains true if the assumption $f(n) = O_L(1)$ is replaced by any Tauberian restriction under which the summability $(A)$ of the series (60) implies its summability $(C, 1)$.

**22.** Inasmuch as ($IX_1$) and ($IX_2$) are not converse mates, it is natural to ask for Tauberian conditions under which the two Eratosthenian summation processes are equivalent (cf. §18). A criterion of this type runs as follows:

(X) *If* $f'(n) = O(1)$, *then the convergence of* $\Sigma f'(n)/n$ *is equivalent to the existence of* $M(f)$.

In fact, the $O$-assumption of (X) implies that of (VII). Hence, one part of (X) follows from (VII). The other part of (X) is contained in ($IX_1$).

Another criterion of the same type as (X) can be formulated as follows:

(XI) *If* $f' \geqq 0$, *then the convergence of* $\Sigma f'(n)/n$ *is equivalent to the existence of* $M(f)$.

In fact, it is clear from (15) that $f \geqq 0$ is implied by $f' \geqq 0$. Hence, (XI) covers a domain common to ($IX_1$) and ($IX_2$).

It will be noted that neither (X) nor (XI) contains (VIII), although (VIII) is a triviality. However, it is seen from the remark made at the end of §21, that the Tauberian assumption of ($IX_2$) could be replaced by one concerning "slow oscillation"; and a corresponding remark holds for the dual theorem, ($IX_1$). In fact, the assumption $f'(n) = O_L(1)$ was used only at the end of the proof of ($IX_1$); its rôle was that of assuring the convergence of the series $\Sigma f'(n)/n$, which was proved to be summable $(A)$ without the assumption $f'(n) = O_L(1)$. Hence, ($IX_1$) remains true if $f'(n) = O_L(1)$ is replaced by any Tauberian condition in virtue of which the summability $(A)$ of $\Sigma f'(n)/n$ implies the convergence of $\Sigma f'(n)/n$.

**23.** Since such a Tauberian condition is supplied by the "high-indices theorem" of Hardy and Littlewood (cf. Ingham [47]), there results one part of the following variant of (X) and (XI):

(XII). *If there exist a sequence* $\{k_m\}$ *and a constant* $\lambda$ *such that*

$$f'(n) = 0 \quad \text{unless} \quad n = k_1, k_2, \cdots, \quad \text{where} \quad k_{m+1}/k_m > \lambda > 1, \tag{65}$$

*then the existence of* $M(f)$ *is equivalent to the convergence of* $\Sigma f'(n)/n$.

The part of (XII) remaining to be proved states that the convergence of $\Sigma f'(n)/n$ implies the existence of $M(f)$, if (65) is satisfied. But (65) implies that, as $n \to \infty$,

$$\sum_{k_m < n} k_m = O(n); \quad \text{whence} \quad \sum_{k_m < n} o(k_m) = o(n) \tag{66}$$

follows by submerging an $\epsilon$ in the usual way. Since the convergence of $\Sigma f'(n)/n$ implies that $f'(n)/n \to 0$ as $n \to \infty$, which in turn implies that $f'(k_m) = o(k_m)$ as $m \to \infty$, it follows that, as $n \to \infty$,

$$(67) \qquad \sum_{k_m < n} |f'(k_m)| = o(n), \quad \text{and so} \quad \sum_{l=1}^{n} |f'(l)| = o(n),$$

by (65). Hence, in order to complete the proof of (XII), it is sufficient to show that, if $\Sigma f'(n)/n$ is convergent and if the last $o$-condition, that is, $|f'(1)| + \cdots + |f'(n)| = o(n)$, is satisfied, then $M(f)$ exists. But this follows from (VIII).

Along the lines of an arithmetical construction of Toeplitz [79], the criterion (XII) was recently proved for the particular case $k_m = p^m$, where $p$ is a fixed prime (van Kampen and Wintner [51]). It now turns out that the true theorem depends only on the lacunary structure of the function $f'(n)$ and has therefore nothing to do with the *arithmetic* assumption of a $p$-adic algorithm.

**24.** The criteria (VII)-(XII) supply a sufficient explanation of the reasons for which the counter-examples establishing (II) and (III) had to be "constructed". It is worth mentioning that there is another, namely function-theoretical, approach to an explanation of this necessity.

Suppose that the signs of absolute value are omitted from the $o$-condition of (VIII). The resulting $o$-condition, which is identical with the tautological assumption of Cesàro (cf. the beginning of §10), can be written in the form $M(f') = 0$. According to (4), this is a necessary condition for the convergence of $\Sigma f'(n)/n$. But if this trivial necessary condition is satisfied, then $\Sigma f'(n)/n$ must be convergent as soon as the function represented by the Dirichlet series $\Sigma f'(n)/n^s$ in the half-plane $\sigma > 1$ has not too weird a behavior near a segment $-\epsilon < t < \epsilon$ of the line $\sigma = 1$. For instance, it is sufficient to assume that the function is bounded for $-\epsilon < t < \epsilon$, , $1 < \sigma < 2$, and that the boundary function on $\sigma = 1$ (which then exists by necessity for almost all $t$ between $t = \pm\epsilon$), is such as to satisfy any of the local conditions assuring the convergence of its Fourier series at $t = 0$ (cf. Karamata [53] and Ingham [46], where certain generalizations of this theorem of M. Riesz are proved).

As to the dual problem, it is clear from (25) that, if the preceding boundedness condition is satisfied, and if $\Sigma f'(n)n^{-s}$ and its boundary function $(-\epsilon < t < \epsilon)$ form a function continuous (and distinct from 0) at $s = 1$, then, since $\zeta(1 + it) \neq 0$, there exists a constant for which $\Sigma f(n)n^{-s} = \text{const.}(s - 1)^{-1} + O(1)$ holds as $s \to 1$, where $\sigma > 1$. It follows therefore from the theorem of Ikehara (cf. §83 below), that, if $\epsilon$ can be chosen arbitrarily large, $M(f)$ must exist as soon as $f(n) = O_L(1)$. Furthermore, this $O_L$-condition can be replaced by a condition of slow oscillation (cf. Karamata [54]).

## Part II

# ARITHMETICAL ALMOST PERIODICITIES



### The Eratosthenian Matrix and Almost Periodic Functions

**25.** For a given $f(n)$, the assignment (15) or (16) defines $f'(n)$ as the function which is transformed into $f(n)$ by the Eratosthenian matrix $(\epsilon_{nm})$. For a given $f(n)$, let $f^*(n)$ denote the function (if any; or, if there are several such functions, then one of them) which is transformed into $f(n)$ by the transposed matrix,

$$(68)\qquad (\epsilon_{nm})^* = (\epsilon_{mn}),$$

of $(\epsilon_{nm})$. Thus (15) is to be replaced by

$$(69)\qquad f(m) = \sum_{n=1}^{\infty} f^*(mn) \equiv \sum_{m\mid n} f^*(n),$$

as seen from the periodic structure of $(\epsilon_{nm})$, described before (19). Correspondingly, the inversion, (24), of (15) becomes replaced by the Möbius inversion

$$(70)\qquad f^*(m) = \sum_{n=1}^{\infty} \mu(n)f(mn) \equiv \sum_{m\mid n} \mu(n/m)f(n),$$

since the transposed matrix of the matrix, (23), of (24) is the (formal) reciprocal of the transposed matrix of $(\epsilon_{nm})$.

Actually, the reciprocal mates (69), (70), obtained by transposing the reciprocal mates (15), (24), present analytical problems not arising for the latter mates. In fact, both linear transformations (15), (24) are defined everywhere and determine each other uniquely. On the other hand, neither of the linear transformations (69), (70) is meaningful without appropriate conditions of convergence. And even if such conditions are satisfied, additional assumptions are needed in order to legitimatize the derangements leading from (69) to its formal inverse (70) or vice versa (cf. §34, §38, §39, §55 below). As a matter of fact, it is a principal characteristic of the ergodic randomness induced on $\mu(n)$ by the distribution of the primes (cf. (iii), §38 below), that the two functions $f$, $f^*$ cannot in general determine each other uniquely, even if they are such as to make both series (69), (70) convergent for every $m$. (Cf. also (i)–(iv), (*1*)–(*4*) in §38– §39).

However, the periodic structure of the columns of the matrix $(\epsilon_{nm})$ makes it seem plausible that, under reasonable restrictions of convergence, the Eratosthenian transformation of $f'(n)$ must generate hidden periodicities in $f(n)$.

These periodicities can beforehand be meant only in the sense, more or less vague, in which the term is used by the geophysicist (who, however, has something very definite, namely the possibility of an anharmonic analysis, in mind), or even only in the sense used by the applied mathematician. But the hidden periodicities may correspond to an almost periodic behavior in the technical sense of the term. The aim of this and of the next chapter is the development of a Fourier theory for the resulting harmonic analysis of the sieve of Eratosthenes.

**26.** For $m = 1, 2, \cdots$ and $n = 1, 2, \cdots$, let

$$e_m(n) = 0 \text{ unless } m \mid n \text{ and } e_m(n) = m \text{ if } m \mid n. \tag{71}$$

Thus, if $m$ is fixed, the function $e_m(n)$ of $n$ has the period $m$, and vanishes except for a single $n$ within a period. Since the exceptional $n$ within a period is characterized by $n \equiv 0 \pmod{m}$, and since the product $e_m(n)e_l(n)$ has the least common multiple, $\{m, l\}$, of $m$ and $l$ as period, it is clear from (2)-(3) that

$$M(e_m e_l) = ml/\{m, l\} = N(e_m e_l) \tag{72}$$

(in fact, the norm $N(g)$ is identical with the mean $M(g)$ if the latter exists and $g \geqq 0$). If the subscript is 1, then (71) is reduced to $e_1(n) = 1$, and therefore (72) to

$$M(e_m) = 1 = N(e_m). \tag{73}$$

It is also clear from (71) that the representation (16) of

$$f(n) = \sum_{d \mid n} f'(d) \tag{74}$$

can be written in the form

$$f(n) = \sum_{m=1}^{\infty} e_m(n) f'(m)/m. \tag{75}$$

Let $f_k(n)$ denote the $k$-th partial sum of the Eratosthenian series (75); so that

$$f_k(n) = \sum_{m=1}^{k} e_m(n) f'(m)/m, \quad \text{i.e.,} \quad f_k(n) = \sum_{d \mid n}^{d \leqq k} f'(d). \tag{76}$$

Since (75) has only a finite number of non-vanishing terms for every fixed $n$,

$$f_k(n) \to f(n) \text{ as } k \to \infty, \qquad (n = 1, 2, \cdots). \tag{77}$$

In view of the periodicity of the functions $e_m(n)$ of $n$, the series (75) suggests a Fourier analysis of the function $f(n)$. However, since *every* $f$, even $f(n) = n!$, has a unique $f'$ by means of which it is representable in the form (75), the Fourier analysis suggested cannot in general exist. On the other hand, it will be easy to find conditions under which the harmonic affinity between (74)-(75) and the sequence of the partial sums (76) is so strong that the ordinary limit relation

(77), which is a triviality (for every fixed $n$), can be replaced by various kinds of convergence in the mean, leading to corresponding classes of almost periodicity.

These classes will be understood to be defined in terms of their appropriate approximability by trigonometric polynomials. For instance, by the almost periodicity ($B$) of a function $f(n)$ will be meant the existence of a sequence of trigonometric polynomials $t_1(n), \cdots, t_k(n), \cdots$ satisfying $N(f - t_k) \to 0$ as $k \to \infty$, where $N(\ )$ is the norm symbol (2); a symbol which is $\bar{M}(\ )$ in the notations of Besicovitch [2]. In this definition, a trigonometric polynomial, $t(n)$, is meant to be any finite sum of the form $a \exp(i\lambda n) + b \exp(i\mu n) + \cdots$, where $a, b, \cdots, \lambda, \mu, \cdots$ are independent of $n$ and $\lambda, \mu, \cdots$ are real.

**27.** Each of the functions (71) of $n$ is periodic (with a period increasing with $m$). Hence, every $e_m = e_m(n)$ is a fixed linear combination of a finite number of roots of unity. In other words, every $e_m(n)$, and therefore every partial sum (76) of (75), is a trigonometric polynomial. This readily leads to the following sufficient criterion:

(XIII) *If the series $\Sigma f'(n)/n$ is absolutely convergent, then the function $f(n)$ is almost periodic* ($B$).

In fact (73), (75), (76) and (2) imply that

$$(78)\qquad N(f - f_k) \leqq \sum_{m=k+1}^{\infty} M(e_m)|f'(m)|/m = \sum_{m=k+1}^{\infty} |f'(m)|/m \to 0$$

as $k \to \infty$, if $\Sigma |f'(m)|/m < \infty$. Hence the truth of (XIII) is clear from the definition of the class ($B$).

This theorem (and the corresponding expansion theorem (XVI), §33, but not the convergence theorem (XVIII), §35) were recently proved with much trouble in the particular case of multiplicative functions $f$ (van Kampen and Wintner [51]). The complications resulted from the use of an arrangement corresponding to the product representation (45) of these particular functions $f$. Such an arrangement, suggested by the whole structure of $f(n)$ in the multiplicative case, does not exist, and cannot therefore become misleading, in the general case.

Since the least common multiple, $\{m, l\}$, of two positive integers cannot exceed their product, $ml$, the assumption in (XIII) is less strict (but, of course, the assertion too is less strict) than in the following theorem:

(XIII bis) *If the double series $\Sigma\Sigma f'(m) f'(l)/\{m, l\}$ is absolutely convergent, then the function $f(n)$ is almost periodic* ($B^2$).

It is understood that the class ($B^2$) results if the requirement $N(f - t_k) \to 0$, which at the end of §26 defined the class $(B) = (B^1)$, is replaced by $N(|f - t_k|^2) \to 0$. Correspondingly, (XIII bis) follows if (73) in the estimate (78) of $N(f - f_k)$ is replaced by (72) in the estimate of $N(|f - f_k|^2)$. In fact, from (75) and (76),

$$|f(n) - f_k(n)|^2 \leqq \sum_{m=k+1}^{\infty} \sum_{l=k+1}^{\infty} e_m(n)e_l(n)|f'(m)f'(l)|/(ml).$$

Although the assumption of (XIII bis) is not a necessary condition for the almost periodicity $(B^2)$ of $f(n)$, the appearance of the double series in (XIII bis) is sufficiently motivated by (21). In this connection, cf. also (XI).

It is clear that similar, though less elegant, criteria follow for any of the classes $(B^q)$ belonging to a fixed Hölder index, $q$. If $f(n) = O(1)$, then the assumption of (XIII) is more inclusive than any of the resulting sufficient conditions. In fact, if $f(n) = O(1)$, then $f(n)$ is either almost periodic $(B^q)$ for every $q$ or is not even almost periodic $(B) = (B^1)$.

**28.** The proof of (XIII) might suggest that the assumption (78) is too drastic, if the assertion is that of (XIII). At any rate, one might expect that the assumptions of (VII), assumptions which, in view of the remark following (VIII), are somewhat milder than what is assumed in (XIII), are sufficient not only for the existence of $M(f)$ but for the almost periodicity $(B)$ of $f$ as well. It is therefore worth proving that such is not the case, not even if the $O(n)$ of (VII) is replaced by the trivial $o(n)$ of (VIII); and that not even the adjunction of the assumption $f'(n) = O(1)$ of (X) is of avail.

(XIV) *The convergence of the series* $\Sigma f'(n)/n$ *and both estimates* $f'(n) = O(1)$, $|f'(1)| + \cdots + |f'(n)| = o(n)$ *together are insufficient for the almost periodicity* $(B)$ *of the function* $f(n)$.

In order to prove this, let $f'(n)$ be so chosen that (i) unless $n$ is a prime, $f'(n) = 0$ and (ii) if $p$ is a prime, $f'(p) = O(1)$ as $p \to \infty$. Since there are only $o(n)$ primes not exceeding $n$, the $o$-assumption of (XIV) is satisfied in virtue of (i) and (ii). Furthermore, (i) and (ii) imply, of course, the $O$-assumption of (XIV). Hence, it suffices to show that (i), (ii) and the convergence of $\Sigma f'(n)/n$, that is, of $\Sigma f'(p)/p$, are insufficient for the almost periodicity $(B)$ of $f(n)$. But it is known that the absolute convergence of $\Sigma f'(p)/p$ is a necessary condition for the almost periodicity $(B)$ of $f(n)$, if (i) is satisfied and $|f'(p)| = 1$ (cf. Hartman and Wintner [42], where necessary and sufficient conditions are given). Hence, it is sufficient to ascertain that $|f'(p)| = 1$ and the convergence of $\Sigma f'(p)/p$ do not imply the absolute convergence of $\Sigma f'(p)/p$. But this is clear from the example $f'(p) = (-1)^j$, where $p$ denotes the $j$-th prime number.

The aim of (XIV) is to show that (XIII) cannot be improved in a certain direction. This does not mean that the convergence of $\Sigma f'(n)/n$ is a necessary condition for the almost periodicity $(B)$ of $f(n)$; cf. (III). All that follows from (VIII)–(XII) is that the almost periodicity $(B)$ of $f(n)$ implies the convergence of $\Sigma f'(n)/n$ under certain Tauberian restrictions. In fact, the existence of $M(f)$ is a necessary condition for the almost periodicity $(B)$ of $f(n)$.

On the other hand, the assumption, $\Sigma |f'(n)|/n < \infty$, of (XIII) implies much more than what is stated by the wording of (XIII). In fact, it implies (78). But (78) states that the *particular* trigonometric polynomials that are supplied by the partial sums, (76), of the Eratosthenian series, (75), of $f(n)$ tend to $f(n)$ in the mean $(B)$. And it will now be shown that there exist functions $f(n)$, of essentially arithmetical significance, which are almost periodic $(B)$ but must be

approximated by a sequence of trigonometric polynomials having a structure substantially distinct from the structure of (76). In other words, the partial sums of the Eratosthenian series of these functions $f(n)$ of class $(B)$ do not have for $f(n)$ the "harmonic affinity" referred to after (77). This implies, of course, that the assumption, $\Sigma \mid f'(n) \mid / n < \infty$, of (XIII) is now violated.

**29.** If $S$ is a set of positive integers, its characteristic function, $f(n) = f_S(n)$, is defined by placing $f(n) = 1$ or $f(n) = 0$ according as $n$ is or is not in $S$. Thus, if $S(n) = f(1) + \cdots + f(n)$, the ratio $S(n)/n$ is the relative frequency of the elements of the set $S$ up to $n$, and the existence of $M(f)$ means that this relative frequency tends to a limit, $M(f)$, as $n \to \infty$. The remark made at the end of §27 is applicable to every characteristic function.

If $S_l$ denotes the set of the integers that are representable as a sum of $l$ squares, then $S_4(n) = n$ and $S_2(n) = o(n)$. Hence the characteristic functions of $S_4$ and $S_2$ are almost periodic $(B)$, having the trivial Fourier expansions $1 + 0 + 0 + \cdots$ and $0 + 0 + \cdots$. It will now be shown that the characteristic function of $S_3$ is almost periodic $(B)$. However, its Fourier expansion, instead of being of a trivial nature, is to be determined from the explicit arithmetical structure of $S_3$ (cf. §45 below).

(XV) *The characteristic function of the set of those integers representable as a sum of three squares is almost periodic* $(B)$ *but such that the partial sums of its Eratosthenian series fail to tend to it in the mean* $(B)$.

If $S$ denotes the set of those positive integers which cannot be represented as a sum of three squares, and if $f(n)$ is the characteristic function of $S$, then the characteristic function to which (XV) refers is the complementary function, $1 - f(n)$. Hence, it is sufficient to prove the assertions of (XV) for $f(n)$.

For a fixed positive integer $k$, let $T_k$ denote the set of those positive integers contained in the progression $4^{k-1}.7, \cdots, 4^{k-1}.(8j - 1), \cdots$. According to a classical result, first proved by Legendre, the set $S$ is identical with $T_1 + T_2 + \cdots$. Since $T_1, T_2, \cdots$ are seen to be mutually disjoint, it follows that $f(n) = g_1(n) + g_2(n) + \cdots$, where $g_k(n)$ denotes the characteristic function of $T_k$. But it is clear from the definition of $T_k$ that the function $g_k(n)$ of $n$ is periodic, having the period $4^{k-1}(8j - 1) - 4^{k-1}(8[j - 1] - 1) = 2^{2k+1}$, and that $g_k(n)$ is 0 except for a single $n$ within a period, for which $n$ it is 1. Hence the function $g_k(n)$ of $n$ has the mean $M(g_k) = 2^{-2k-1}$. But this mean is identical with the norm $N(g_k)$, since $g_k(n) \geqq 0$. It follows therefore from $f(n) = g_1(n) + g_2(n) + \cdots$ that

$$(79) \qquad N\left(f - \sum_{m=1}^{k} g_m\right) \leqq \sum_{m=k+1}^{\infty} M(g_m) = \sum_{m=k+1}^{\infty} 2^{-2m-1} \to 0$$

as $k \to \infty$. Since every $g_m(n)$ is periodic, (79) implies that $f(n)$ is almost periodic $(B)$.

The remaining statement of (XV) is that (79) cannot be replaced by (78); cf. the comments made at the end of §28. But the failure of the limit relation

(78) is quite obvious from (74)-(76) and from the structure of the functions $g_k(n)$ corresponding to the decomposition $S = T_1 + T_2 + \cdots$. An equivalent, though indirect, proof of this failure will follow by comparing ($80_3$) and (81) with (97) below.

It may be mentioned that a theorem of Landau [57], according to which the asymptotic relative frequency ("density") of the integers contained in $S_3$ exists and equals 5/6, is a corollary of (79). In fact, since every $g_m(n)$ is periodic, (79) implies that $M(f)$ exists and is represented by

$$M(f) = \sum_{m=1}^{\infty} M(g_m) = \sum_{m=1}^{\infty} 2^{-2m-1} = 1/6. \tag{79 bis}$$

**30.** Inasmuch as only the periodic character of the mutually disjoint sets $T_1, T_2, \cdots$ appears to have been used, one might expect that the above proof for the almost periodicity ($B$) of $f(n)$ could be transcribed so as to hold for the characteristic function of any set $S$ having the following property: There exists a sequence of positive integers $d_1, d_2, \cdots$ such that an $n$ is or is not in $S$ according as $d_k \mid n$ does or does not hold for at least one $k$ (clearly, it can be assumed without loss of generality* that $d_k \mid d_i$ only when $k = i$). If, corresponding to a set $S$, there exists such a sequence $\{d_k\}$, let $S$ be called a division set (belonging to the sequence $\{d_k\}$, which can be chosen arbitrarily).

Besicovitch [3] constructed a sequence $\{d_k\}$ such that, if $f(n)$ denotes the characteristic function of the corresponding division set $S$, then $M(f)$ does not exist. Since the existence of $M(f)$ is a necessary condition for the almost periodicity ($B$) of any function, it follows that there exist division sets $S$ for which $f(n) = f_S(n)$ is not almost periodic ($B$). However, something can be saved from §29.

(XV*) *If $S$ is a division set, and if $f(n)$ denotes its characteristic function, then the existence of $M(f)$ is sufficient (and, of course, necessary) for the almost periodicity ($B$) of $f(n)$.*

The assumption that the set $S$ be a division set is essential for the truth of (XV*). In fact, if the set $S$ is unrestricted, it can obviously be so chosen that, if $f_e(n)$, $f_o(n)$ and $f(n)$ respectively denote the characteristic functions of the sets consisting of all even, of all odd, and of all, integers contained in $S$, then $M(f)$ does, but $M(f_e)$ does not, exist. Then it is clear from $f_e(n) + f_o(n) = f(n)$ that $M(f_o)$, and therefore $M(f_o - f_e)$, does not exist. It follows therefore from the obvious identity $f_e(n) - f_o(n) = (-1)^n f(n)$, that the Fourier constant $M(f^\lambda)$, where $f^\lambda(n) = f(n)\exp(-2\pi i\lambda n)$, fails to exist for $\lambda = \frac{1}{2}$. This implies, of course, that $f(n)$ cannot be almost periodic ($B$). Nevertheless, $f(n)$ is a characteristic function for which $M(f)$ exists.

* On the other hand, the assumption that $(d_k, d_i) = 1$ unless $k = i$ would involve a rather serious loss of generality. In fact, if any two of the integers $d_1, d_2, \cdots$ are coprime, it is clear that the set $S$ belonging to the sequence $d_1, d_2, \cdots$ is multiplicative (that is, that the characteristic function of $S$ is a multiplicative function in the sense of §13), and is, as a matter of fact, a multiplicative set of particular structure. As to arbitrary multiplicative sets, cf. §52 below.

In order to prove (XV*), let $R_k$ denote the set containing an $n$ if and only if $d_i \mid n$ holds for at least one $i$ not exceeding $k$. Then $R_k$ is a subset of $R_{k+1}$, and $R_k \to S$ as $k \to \infty$. In other words, if $h_k(n)$ denotes the characteristic function of $R_k$, then $h_k(n) \leqq h_{k+1}(n)$ and $h_k(n) \to f(n)$ as $k \to \infty$ hold for every fixed $n$; so that $h_k(n) \leqq f(n)$. It is also seen that the function $h_k(n)$ of $n$ is periodic; in fact, a Euclidean algorithm shows that the least common multiple of $d_1, \cdots, d_k$ is a period. In particular, the mean $M(h_k)$ exists for every $k$. Furthermore, $M(h_1) \leqq M(h_2) \leqq \cdots \leqq 1$, since $h_1(n) \leqq h_2(n) \leqq \cdots \leqq 1$ for every $n$. Hence, $M(h_k)$ tends to a limit as $k \to \infty$. According to Davenport and Erdös [12], the latter limit has the value

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} f(m), \quad \text{even if} \quad M(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} f(m)$$

does not exist (as it does not in the example of Besicovitch, referred to before). In particular, if $M(f)$ exists, then the limit of $M(h_k)$ as $k \to \infty$ must be $M(f)$. Since this can be written in the form $M(f - h_k) \to 0$, it follows from $h_k(n) \leqq f(n)$ that $M(|f - h_k|) \to 0$. Hence, (XV*) is implied by the periodicity of the functions $h_k(n)$.

### The Transposed Matrix and Fourier Series

**31.** The preceding chapter is only preparatory to the Fourier theory announced at the end of §25. This theory will now be developed.

The Fourier analysis will automatically lead to the transposed matrix, (68), of the Eratosthenian matrix. In fact, the Fourier constant of $f(n)$ will appear in the form (69), if the operator * is applied to the function $f'(n)/n$. However, the occurrence of the Möbius function in the inversion, (70), of (69) will induce phenomena of divergence corresponding to those in the theory of ordinary Fourier series.

**32.** It is clear from (71) that the matrix elements of the Eratosthenian transformation (74)-(75) can be written in the form

$$(80_1) \qquad e_m(n) = \sum_{j=1}^{m} \exp\,(2\pi i n j/m) = \begin{cases} m \text{ if } n = m, 2m, \cdots, \\ 0 \text{ unless } m \mid n, \end{cases}$$

and so

$$(80_2) \qquad e_1(n) = 1 \quad \text{and} \quad e_{p^k}(n) = \begin{cases} p^k \text{ if } k \leqq a, \\ 0 \text{ if } k > a, \end{cases} \qquad n = p^a q^b \cdots,$$

where $a, b, \cdots$ are non-negative integers and $p, q, \cdots$ denote distinct primes.

Let $c_m(n)$ be defined as the expression obtained by retaining in the sum $(80_1)$ only those terms, $\varphi(m)$ in number, for which $j$ is relatively prime to $m$. Thus

$$(80_3) \qquad c_m(n) = \sum_{l}{}' \exp(2\pi i n l/m), \quad \text{where } 1 \leqq l \leqq m \text{ and } (l, m) = 1.$$

Then $e_m(n) = \sum_{d\mid m} c_d(n)$, by $(80_1)$. Hence, from (15),

$$(80_4) \qquad f'(m) = c_m(n) \quad \text{if} \quad f(m) = e_m(n) \quad \text{for fixed } n.$$

Consequently, from (18) and (17),

$$(80_5) \qquad c_1(n) = 1 \quad \text{and} \quad c_{p^k}(n) = e_{p^k}(n) - e_{p^{k-1}}(n), \qquad (k = 1, 2, \cdots),$$

since $e_1(n) = 1$, by $(80_1)$. Also

$$(80_6) \qquad c_m(1) = \mu(m)$$

in virtue of the definition (22). Furthermore, from $(80_5)$ and $(80_2)$,

$$(80_7) \qquad c_p(n) = -1 \quad \text{unless} \quad p \mid n \quad \text{and} \quad c_p(n) = p - 1 \quad \text{if } p \mid n.$$

It is clear from (71) that, if $n$ is fixed, $e_m(n)$ is a multiplicative function of $m$. It follows therefore from $(80_4)$ and (39) that the same is true of $c_m(n)$. This proves without any calculation the principal formal property of the so-called Ramanujan sums, $(80_3)$.

Even the "explicit" representation

$$(80_8) \qquad c_m(n)/\phi(m) = \mu(m/(m, n))/\phi(m/(m, n)),$$

pointed out by O. Hölder, follows with a minimum of labor, if it is observed that, in virtue of $(80_4)$, (71) and (24),

$$(80_9) \qquad c_m(n) = \sum_{d\mid(m,n)} \mu(m/d)d,$$

where $(m, n)$ denotes the greatest common divisor of $m$ and $n$.

**33.** If $(80_1)$ is combined with the Eratosthenian identity (75), it is easy to show that the assertion of (XIII) can be completed as follows:

(XVI) *If the series $\Sigma f'(n)/n$ is absolutely convergent, then the function $f(n)$ is almost periodic $(B)$ and has the Fourier series $(B)$*

$$(81) \qquad f(n) \sim \sum_{m=1}^{\infty} a_m c_m(n),$$

*where $c_m(n)$ is the Ramanujan sum $(80_3)$ and the Fourier coefficients $a_m$ are given by*

$$(82) \qquad a_m = \sum_{m\mid n} \frac{f'(n)}{n}; \qquad (a_1 = M(f) = \sum f'(n)/n).$$

In fact, insertion of $(80_1)$ into (76) gives

$$f_k(n) = \sum_{m=1}^{k} \frac{f'(m)}{m} \sum_{j=1}^{m} \exp\,(2\pi inj/m).$$

Since the absolute convergence of $\Sigma f'(n)/n$ implies the relation (78), according to which $f_k(n)$ tends to $f(n)$ in the mean of the space $(B)$, it follows, by letting $k \to \infty$, that

$$(81 \text{ bis}) \qquad f(n) \sim \sum_{h=1}^{\infty} \sum_{\lambda=r/s} \frac{f'(hs)}{hs} \exp(2\pi inr/s),$$

where $(r, s) = 1$, and $\lambda$ ranges over all rational numbers $r/s$ contained in the interval $0 < \lambda \leqq 1$. Since the definitions $(80_3)$, (82) imply that (81 bis) is identical with (81), the proof of (XVI) is complete.

**34.** Inasmuch as (75) is an identity, one might expect that the representation of $f(n)$ in terms of its Fourier expansion (81)-(82) also is an identical relation. But such an expectation is erroneous.

(XVII) *The absolute convergence of* $\Sigma f'(n)/n$, *which according to* (XVI) *implies for* $f(n)$ *the Fourier series* (81), *is compatible with the divergence of the latter.*

In fact, it is clear from $(80_6)$ and (82) that the $k$-th partial sum of the Fourier series (81) at $n = 1$ is

$$(83_0) \qquad s_k = \sum_{m=1}^{k} \mu(m) \sum_{m \mid n} t_n\,,$$

if $t_n$ is an abbreviation for $f'(n)/n$. Hence, in order to prove (XVII), it suffices to show the existence of sequences $\{t_n\}$ for which the series $\Sigma\, t_n$ is absolutely convergent but the corresponding sequence $\{s_k\}$, defined by $(83_0)$, is divergent. Consequently, it is more than sufficient to show that $\Sigma \mid t_n \mid < \infty$ is compatible with $s_k \neq O(1)$ as $k \to \infty$. But $(83_0)$ can be written in the form

$$(83_1) \qquad s_n = \sum_{m=1}^{\infty} \alpha_{nm} t_m\,; \qquad\qquad (83_2) \qquad \alpha_{nm} = \sum_{d \mid m}^{d \leqq n} \mu(d)$$

(if $k$, $m$, $n$ in $(83_0)$ are replaced by $n$, $d$, $m$ respectively). Furthermore, it is clear from the Lebesgue-Toeplitz construction that, if a linear transformation $(83_1)$ violates the norm condition

$$\limsup_{n\to\infty,\, m\to\infty} \mid \alpha_{nm} \mid < \infty,$$

then it certainly cannot transform *every* sequence $\{t_n\}$ satisfying $\Sigma \mid t_n \mid < \infty$ into a sequence $\{s_n\}$ satisfying $s_n = O(1)$ as $n \to \infty$. Consequently, it is sufficient to show that the requirement expressed by the last formula line is violated, if the matrix $(\alpha_{nm})$ is defined by $(83_2)$. In other words, it is sufficient to ascertain the existence of a function pair $n = n(k)$, $m = m(k)$ satisfying

$$(84) \qquad n(k) \to \infty, \qquad m(k) \to \infty \quad \text{and} \quad \left| \sum_{d \mid m(k)}^{d \leqq n(k)} \mu(d) \right| \to \infty \quad \text{as} \quad k \to \infty.$$

But such a function pair can readily be constructed by using products of lacunary sequences of primes.

The construction necessarily leads to a *steep* "curve" $n = n(k)$, $m = m(k)$ in the $(n, m)$-"plane", since, according to $(83_2)$,

$$(85) \qquad \mid \alpha_{nm} \mid \leqq n \quad \text{and} \quad \mid \alpha_{nm} \mid \leqq d(m),$$

if $d(m)$ denotes the number of the divisors of $m$ (in particular, every column as well as every row of the matrix $(\alpha_{nm})$ consists of a bounded sequence of numbers).

The expressions considered above were the partial sums $(83_0)$ of the Fourier series (81) at $n = 1$. It is clear from $(80_8)$ that the construction could be carried

out for every fixed value of $n$ in the Fourier series (81). In view of the diagonal principle of Cantor, it might even be possible to construct an example satisfying the assumption of (XVI) and rendering the Fourier series (81) divergent for all $n$.

If $m \leqq n$, then $d \mid m$ implies that $d \leqq n$, and so it is clear from $(83_2)$ and (24 bis) that

$$\alpha_{n1} = 1 \quad \text{and} \quad \alpha_{nm} = 0 \quad \text{for} \quad 2 \leqq m \leqq n. \tag{85 bis}$$

But if $n < m$, the value of the sum $(83_2)$ is not under explicit control. Thus the fact that the full force of the assumption $\Sigma \mid t_n \mid < \infty$ does not appear to have been used in the reduction of $s_n \neq O(1)$ to (84), is suggested by an observation of Hadamard [33] concerning the converse of the Dirichlet-Dedekind criteria for convergent partial summation.

**35.** In view of (XVI) and (XVII), the situation is comparable to that in the theory of ordinary Fourier series of functions $f(t)$ on the interval $0 \leqq t < 2\pi$ (the common assumption, $\Sigma \mid f'(n) \mid /n < \infty$, of (XVI) and (XVII) corresponds, besides to $L$-integrability, to continuity or integrated continuity near every $t$). Correspondingly, there arises the need for a criterion similar to the convergence theorem of Dirichlet-Jordan, for instance. It is indicated by (85) that a sufficient criterion of this type will somehow involve the function $d(m)$ of $m$. Actually, a simple (and, as will be seen in §36 and §46–§54, quite useful) analogue of the Dirichlet-Jordan theorem turns out to be the following:

(XVIII) *If the series* $\Sigma\, d(n)f'(n)/n$ *is absolutely convergent, then not only* (81) *holds for* (82) *but also*

$$f(n) = \sum_{m=1}^{\infty} a_m c_m(n), \qquad (n = 1, 2, \cdots). \tag{86}$$

*Furthermore, although* $f(n) \neq O(1)$ *in general, the convergence of the Fourier series* (86) *is absolute (but, if* $f(n) \neq O(1)$, *certainly not uniform) for all* $n$.

The assumption of (XVIII) is more severe than that of (XVI)–(XVII), but only slightly so, since

$$1 \leqq d(n) \neq O(\log^{1/\epsilon} n) \quad \text{if} \quad \epsilon > 0, \quad \text{but} \quad d(n) = O(n^{1/\log\log n}).$$

It is clear from $(80_9)$ that $\mid c_m(n) \mid \leqq \sigma\,((m, n))$, if $\sigma(k)$ is the sum of the divisors of $k$. Hence, if $\sigma_n$ denotes the maximum of $\sigma(k)$ for $1 \leqq k \leqq n$, then $\mid c_m(n) \mid \leqq \sigma_n$. Since $\sigma_n$ is independent of $m$, it follows from (82) that the series (86) belonging to an arbitrary $n$ is absolutely convergent if

$$\sum_{m=1}^{\infty} \left| \sum_{m \mid n} f'(n)/n \right| \leqq \sum_{m=1}^{\infty} \sum_{m \mid n} \mid f'(n)/n \mid < \infty. \tag{87}$$

Since the range of the double summation contains a fixed positive integer, say $k$, exactly $d(k)$ times, the last double series can be contracted into $\Sigma\, d(k) \mid f'(k) \mid /k$. This proves that part of (XVIII) asserting the absolute convergence of the

series (86). That the sum of the series (86) is precisely $f(n)$, is clear from the proof of (XVI), since the convergence of the double series on the right of (87) ensures the legitimacy of the *formal rearrangement* of (75) into (86).

**36.** Since the sum $(80_3)$ is a periodic, and therefore bounded, function of $n$ for every $m$, the parenthetical remark of (XVIII) is obvious. Hence, in order to complete the proof of (XVIII), it is sufficient to exhibit functions $f(n)$ satisfying the assumption of (XVIII) but violating $f(n) = O(1)$.

An interesting function having these properties is $f(n) = \omega(n) - \nu(n)$, where $\omega(n)$ denotes the number of all, and $\nu(n)$ the number of all distinct, prime divisors of $n$. In other words, $\omega(n)$ is the sum of the exponents $\alpha, \beta, \cdots$ in the factorization $p^\alpha q^\beta \cdots$ of $n$ into powers of $\nu(n)$ distinct primes $p, q, \cdots$, where it is understood that $\omega(1) = 0$ and $\nu(1) = 0$. Thus it is seen from (15) that the function $\nu'(n)$ belonging to $\nu(n)$ is 1 or 0 according as $n$ is or is not a prime. Correspondingly, $\omega'(n)$ is 1 or 0 according as $n$ is or is not a prime power (which can be a prime).

Since the connection (15) between the functions $f(n), f'(n)$ is distributive, it follows by subtraction that, if $f(n) = \omega(n) - \nu(n)$,

$$f'(n) = 1 \quad \text{if} \quad n = p^2, p^3, \cdots \quad \text{and} \quad f'(n) = 0 \quad \text{if} \quad n \neq p^2, p^3, \cdots,$$

where $p$ denotes any prime number. Inasmuch as the number, $d(n)$, of all divisors of $n$ is $k + 1$ when $n = p^k$, it follows that

$$\sum_{n=1}^{\infty} d(n) \, |f'(n)|/n = \sum_{k=2}^{\infty} \sum_{p} (k + 1)/p^k = \sum_{k=2}^{\infty} (k + 1) \, O(2^{-k}) < \infty.$$

Hence, the assumption of (XVIII) is satisfied by $f(n) = \omega(n) - \nu(n)$, although $f(n) \neq O(1)$.

It also follows that the series (82) is the sum of the reciprocal values of those prime powers (distinct from $p^0 = 1$) which are multiples of $m$ but are not primes. This implies that the Fourier constant $a_m$ vanishes unless $m = 1, p, p^2, \cdots$, and that $a_1 = \Sigma f'(n)/n = M(f)$ has the value $\Sigma \, (p^2 - p)^{-1}$, where $p$ ranges over all primes.

**37.** Since two trigonometric polynomials $(80_3)$ belonging to two distinct values of $m$ have no frequency in common, it is clear that, if a series (86), where $a_1$, $a_2, \cdots$ are arbitrary constants, is uniformly convergent for all $n$, and if $f(n)$ is the function defined by (86), then $f(n)$ is uniformly almost periodic (that is, almost periodic in the sense of Bohr) and has the Fourier series (81).

However, it does not follow that the Fourier constant $a_m$ must then be representable in the form (82). Furthermore, an adaptation of standard constructions leads to a uniformly almost periodic function $f(n)$ having a Fourier series of the form (81) which, however, fails to converge uniformly. Finally, the uniformly almost periodic function $f(n) = \exp(2\pi i n/3)$, being identical with its Fourier series, has a Fourier series not representable in the form (81); in fact, the sum $(80_3)$ belonging to $m = 3$ consists of $\varphi(3) > 1$ terms.

The transcription of the proof, (78), of (XIII) to the case of uniform almost periodicity leads only to the sufficient criterion represented by

$$\sum |f'(n)| < \infty. \tag{88}$$

In fact, since the function (71) of $n$ is periodic and has an absolute value not exceeding $m$, the Eratosthenian series (75) is a uniformly convergent series of periodic functions, if (88) is satisfied.

Clearly, (88) entails the assumption of (XVIII). Actually, the Fourier series (86) of $f(n)$ is absolutely-uniformly convergent, if (88) is satisfied.

In fact, since the sum ($80_3$) consists of $\varphi(m)$ terms of absolute value 1, it is clear from (82) that the series (86) is absolutely-uniformly convergent whenever

$$\sum_{m=1}^{\infty} \phi(m) \left| \sum_{m \mid n} f'(n)/n \right| \leqq \sum_{m=1}^{\infty} \sum_{m \mid n} \phi(m) \, |f'(n)/n| < \infty. \tag{87*}$$

But the range of the double summation contains a fixed integer, say $k$, exactly $d(k)$ times, and the corresponding $f'(n)/n$ each time is multiplied by $\varphi(d)$, where $d$ runs through the $d(k)$ divisors of $k$. Hence, the requirement expressed by the convergence of the double series (on the right) is identical with

$$\sum_{k=1}^{\infty} |f'(k)/k| \sum_{d \mid k} \phi(d) < \infty \quad \text{or, since} \quad \sum_{d \mid k} \phi(d) = k, \tag{88 bis}$$

with the assumption (88). This proves the assertion.

It does not follow that the assumption (88) implies for the Fourier series (86) a convergent majorant $\Sigma\, \alpha_m$, where $\alpha_m (> 0)$ is independent of $n$. Nevertheless, it is clear from the preceding remarks that condition (88) as a sufficient criterion for the uniform almost periodicity of $f(n)$ is of a trivial nature. On the other hand, if (88) is replaced by the assumption that $\Sigma\, f'(n)$ is a convergent series (an assumption under which ($29_3$) has been established and which therefore is more than sufficient for the existence of $M(f)$ and for the convergence of $\Sigma\, f'(n)/n$), then $f(n)$ need not be uniformly almost periodic; not even if $f(n)$ is assumed to be almost periodic ($B$). In fact, if $f(n) = \phi(n)/n$, then $f(n)$ is almost periodic ($B$) but is not uniformly almost periodic (cf. §48 and §49 below), although $\Sigma\, f'(n)$ is convergent, since (15) holds for $f(n) = \phi(n)/n$ if $f'(n) = \mu(n)/n$ (the convergence of $\Sigma\, \mu(n)/n$ is assured by the prime number theorem). Conversely, the convergence of $\Sigma\, f'(n)$ is easily seen to be unnecessary for the uniform almost periodicity of $f(n)$.

All of this shows that a necessary and sufficient condition characterizing those functions $f'(n)$ for which the function $f(n)$ defined by (15) becomes uniformly almost periodic must involve *arithmetical* criteria of a high degree of delicacy. But the existence of a reasonably explicit condition of this type is problematic; the more so as, from the point of view of the standard arithmetical functions, the notion of uniform almost periodicity appears to be a somewhat artificial construction* (cf. §49–§51 below).

* It is interesting that in this regard the situation is the same as in the applications of the notion of uniform almost periodicity to general dynamical problems. Cf. A. Wintner, Proc. Nat. Acad. Sci., vol. 27 (1941), pp. 311–314; N. Wiener and A. Wintner, Amer. Journ. of Math., vol. 63 (1941), pp. 794–824; P. Hartman and A. Wintner, ibid., vol. 65 (1943).

## The Sieve of Eratosthenes and Harmonic Analysis

**38.** The connection between the assumptions and assertions of (XIV)–(XVIII), on the one hand, and the fact that (75) is an identity for every $f$, on the other hand, becomes revealing if the details of the formal rearrangement, mentioned at the end of §35, are followed in case of a few functions of arithmetical interest.

(i) Let either $f(n) = \nu(n)$ or $f(n) = \omega(n)$, where $\nu(n)$, $\omega(n)$ are the functions defined at the beginning of §36. Both functions (32) are of the order of $\log\log n$ in either case, since

$$\sum_{p<n} \frac{1}{p} \sim \log\log n, \qquad \sum_p \sum_{k=2} \frac{1}{p^k} < \infty, \qquad \frac{1}{n}\sum_{m=1}^{n} \nu(m) \sim \log\log n$$

and $M(\omega - \nu) = \Sigma\,(p^2 - p)^{-1}$, by the end of §36. In particular, neither $M(f)$ exists. However, since $\nu(n)$ is the number of all primes $p$ satisfying $p \mid n$, and since $(80_7)$ states that $1 + c_p(n)$ is $p$ or 0 according as $p$ does or does not satisfy $p \mid n$, it is clear that

$$\nu(n) = \sum_p \frac{1 + c_p(n)}{p}$$

is an identity in $n$. This infinite series has only a finite number of non-vanishing terms for every fixed $n$. It can be thought of as representing the series (86) in the case $f = \nu$; in the sense that, in accordance with the fact that both functions (32) belonging to $f = \nu$ are asymptotically equal to $\log\log n$, the coefficient $a_1 = M(f)$ turns out to be $\Sigma\, p^{-1} = \infty$. If $m > 1$, the identification of the last formula line with (86) gives $a_m = 0$ unless $m = p$, and $a_p = p^{-1}$ for every $p$. And all of this is precisely what is assigned by (82), even if $m = 1$ is not excluded. In fact, it was seen in §36 that $\nu'(n) = 0$ unless $n = p$ and that $\nu'(p) = 1$ for every $p$.

(ii) It is clear from the results of §36 that the preceding remarks remain valid if the number, $\nu(n)$, of all distinct prime divisors of $n$ is replaced by the number, $\omega(n)$, of all prime divisors of $n$. On the other hand, the situation becomes almost the opposite if either of these functions $f(n)$ is replaced by $f(n) = d(n)$, the number of all divisors of $n$. In fact, Ramanujan ([69], p. 186) has shown (on the basis of the prime number theorem) that the trigonometric series (86) having the coefficient $a_m = -m^{-1} \log m$ for $m = 1, 2, \cdots$ converges, and has the sum $d(n)$, for every $n$. Since $d'(n) = 1$ for every $n$ in virtue of (15), and since $d(1) + \cdots + d(n) \sim n \log n$, both functions (32) belonging to $f(n) = d(n)$ are of the order of $\log n$. In particular, $M(f)$ must be identified with $\infty$. But $a_m = -m^{-1} \log m$ is 0 when $m = 1$, and so the identification $a_1 = M(f)$, which for $f = \nu$ gave $\infty = \infty$, now gives $0 = \infty$. Finally, since $f'(n) = 1$ for every $n$, the series (82) gives $-m^{-1} \log m = \infty$ for $m = 1, 2, \cdots$. Nevertheless, the series (86), which for $f = \nu$ was a trigonometric series only in a metaphorical sense, is now *bona fide* trigonometric.

(iii) It may be shown that neither of the series just mentioned is a Fourier series in any sense (for instance, in the sense to be used in (*), §55 below). A

function with a Fourier series $(B)$ will now be considered. First, the prime number theorem (for the arithmetic progressions) is known to imply that each of the Kluyver series $\Sigma\, \mu(n)/n$, $\Sigma\, \mu(2n)/n$, $\cdots$, $\Sigma\, \mu(mn)/n$, $\cdots$ is convergent and has the sum 0, which means that each of the series (82) converges to 0 when $f'(n) = \mu(n)$. According to (22), the corresponding $f(n)$ is 0 unless $n = 1$, which implies that $f(n)$ is almost periodic $(B)$, with $f(n) \sim 0 + 0 + \cdots$ as Fourier series. Consequently, the assertions of (XVI) are true, although the assumption, $\Sigma \mid \mu(n) \mid /n < \infty$, of (XVI) is violated. Furthermore, (86) becomes $1 = 0 + 0 + \cdots$ when $n = 1$ but is correct for every $n > 1$.

(iv) A dual of the example (iii) results if $a_m$ is assigned to be $m^{-1}$ for every $m$ and the function $f'$ is then defined by the formal inversion of (82). The formal inversion of (69) being (70), the value of $f'(m)$ is $\Sigma\, \mu(n)(mn)^{-1}$, i.e., $m^{-1}\, \Sigma\, \mu(n)/n$, which, by the prime number theorem, is 0 for every $m$. Correspondingly, $(80_6)$ shows that the sum of the series (86) belonging to $a_m = m^{-1}$ is 0 when $n = 1$. According to Ramanujan ([69], p. 185), this implies that the sum of the series (86) belonging to $a_m = m^{-1}$ is 0 for every $n$. Since $f'(m) = 0$ for every $m$, it follows from (15) that (86) is true for every $n$.

**39.** These examples suggest the following observations:

(*1*) Although $f(n)$ was almost periodic $(B)$ in (iii), the convergence of the series (82) depended on the prime number theorem. Correspondingly, the series (82) need not converge whenever $f(n)$ is almost periodic $(B)$. Examples to this effect can be constructed by observing that, if $e(n)$ denotes the simple vibration $e^{i\lambda n}$ belonging to a fixed real $\lambda$, then $e'(n) = \sum_{d \mid n} \mu(n/d) e^{i\lambda d}$, by (24).

(*2*) There exist sequences $\alpha_1, \alpha_2, \cdots$ for which each of the series $\Sigma\, \alpha_n$, $\Sigma\, \alpha_{2n}, \cdots, \Sigma\, \alpha_{mn}, \cdots$ is convergent and has the sum 1 (for a class of simple examples, derived by ordinary harmonic analysis, cf. Rajchman [68]). Hence, if $f'(n) = n\alpha_n$, then each of the series (82) is convergent and $a_1 = a_2 = \cdots = 1$ holds for the corresponding $f(n)$. This means that the formal rearrangements referred to at the beginning of §38 can lead, instead of to (81)-(82), to (82) and to the trigonometric series $\Sigma\, c_m(n)$; a series which corresponds to $\frac{1}{2} + \Sigma \cos mt$ in the analogy pointed out before (XVIII).

(*3*) Let $\beta_1, \beta_2, \cdots$ be a sequence having the same property as, but distinct from, the sequence $\alpha_1, \alpha_2, \cdots$ used before. Then each of the series (82) belonging to $f'(n) = n(\alpha_n - \beta_n)$ converges and has the sum 0, as in (iv). However, the situation is not the same as in (iv). In fact, since $\alpha_n \neq \beta_n$ for at least one $n$, it follows from (15) that the function $f(n)$ does not, although the coefficient $a_m$ of the series (86) does, vanish identically; so that the relation (86), which in (iv) was true (for every $n$), is now false (for some $n$).

(*4*) The preceding example and (iv) imply that Cantor's uniqueness theorem can in no sense be paralleled in the analogy referred to before (XVIII).

(*5*) In view of the problem of Fourier constants of periodic functions of class $(L)$, there cannot be expected a direct characterization of sequences $a_1, a_2, \cdots$

for which the trigonometric series $\Sigma\, a_m c_m(n)$ is the Fourier series of some $f(n)$ of class $(B)$. On the other hand, since $(80_3)$ consists of $\phi(m)$ terms, it is clear from the analogue of the Fischer-Riesz theorem that $\Sigma\, a_m c_m(n)$ is the Fourier series of some $f(n)$ of class $(B^2)$ if and only if $\Sigma\, \phi(m) \mid a_m \mid^2 < \infty$. Cf., however, (XV) and the remark made at the end of §27.

**40.** For an arbitrary function $f(n)$, let

(89) $$f_\lambda(n) = f(n) \exp(-2\pi i \lambda n),$$

where $\lambda$ is a real number. Since $\lambda$ can be reduced mod 1, it will always be assumed that either $0 < \lambda \leqq 1$ or $0 \leqq \lambda < 1$. If the mean, (3), of (89) exists for a fixed $\lambda$, the value $M(f_\lambda)$ is called the $\lambda$-th amplitude of the function $f(n)$. In particular, the 0-th amplitude is $M(f)$, if $M(f)$ exists.

The function $f(n)$ is said to have an amplitude function, $M(f_\lambda)$, $0 \leqq \lambda < 1$, if its $\lambda$-th amplitude exists for every $\lambda$. In this and only in this case has $f(n)$ a Fourier series

(90) $$f(n) \sim \sum_k M(f_{\lambda_k}) \exp(2\pi i \lambda_k n),$$

if (90) is defined to be just an abbreviation for the following pair of statements:

($i$) The mean $M(f_\lambda)$ of (89) exists for every $\lambda$ and
($ii$) the value of $M(f_\lambda)$ is 0 unless $\lambda$ is a $\lambda_k$ (mod 1).

If $f(n)$ is almost periodic $(B)$ and (90) is its Fourier series $(B)$, then the latter is the Fourier series in the sense ($i$)-($ii$) also. The converse is true provided that $f(n)$ is supposed to be almost periodic $(B)$. The latter proviso cannot be omitted. In other words, the existence of an amplitude function is necessary but not sufficient for the almost periodicity $(B)$ of $f(n)$.

For instance, let $f(n)$ be the characteristic function of a set $S$ (as defined at the beginning of §29), and suppose that $f(n)$ is almost periodic $(B)$. Then it is almost periodic $(B^2)$, since $f(n) = O(1)$. Consequently, by Parseval's relation.

(91) $$M(f) - M(f)^2 = \sum_{0<\lambda<1} \mid M(f_\lambda) \mid^2,$$

since $f = f^2$ and $M(f_0) = M(f)$. But the harmonic law of large numbers implies that if a set $S$ of positive integers is chosen "at random", then its characteristic function, $f(n)$, possesses an amplitude function, $M(f_\lambda)$, which vanishes for $0 < \lambda < 1$ but is $\frac{1}{2}$ for $\lambda = 0$ (Wiener and Wintner [86]). And such an $f(n)$ cannot be almost periodic $(B)$, since it reduces (91) to $\frac{1}{2} - (\frac{1}{2})^2 = 0$.

It should be mentioned that, $k$ being an integer, the notation (90) involves a tacit assumption. No such assumption is needed if $f(n)$ is either almost periodic $(B)$ or such that $g = f^2$ has a finite norm (2). In fact, it is clear in the first case, and it follows from an adaptation of Bessel's inequality in the second case, that $M(f_\lambda) = 0$ for every $\lambda$ not contained in an enumerable set. But if $f(n)$ is not

almost periodic $(B)$ and the assumption $N(f^2) < \infty$ is lightened to $N(f) < \infty$ (or, for that matter, to $N(f^{2-\epsilon}) < \infty$), it has never been proved or disproved that the amplitude function must vanish on the complement of an enumerable $\lambda$-set. On the other hand, not even $N(f) < \infty$ is assumed in the definition $(i)$-$(ii)$ following (90). Cf. Wintner [91].

**41.** If an $f(n)$ is such that, for a fixed $l$, the function

$$f^{(l)}(n) = f(n)\bar{f}(n + l) \tag{92}$$

of $n$ possesses a mean, $M(f^{(l)})$, the latter is called the $l$-th auto-correlation coefficient of $f(n)$. In particular, if $M(|f|^2)$ exists, it is the 0-th coefficient of auto-correlation, provided that $l = 0$ is allowed in the preceding definition. In view of (130), §59 and (**), §60 below, it will be convenient to assume that $l$ denotes a *positive* integer. This agreement deviates from the usual terminology, according to which $f(n)$ is said to possess an auto-correlation function if $M(f^{(l)})$ exists for every *non-negative* integer $l$; a terminology which alone makes possible a spectral characterization of the class $(B^2)$ of almost periodic functions (cf. Wiener and Wintner, [87], Theorem 3).

If $f(n)$ has an amplitude function, that is, a Fourier series in the sense defined in §40, then substitution of (90) into the product (92) gives

$$M(f^{(l)}) = \sum_k |M(f_{\lambda_k})|^2 \exp(2\pi i \lambda_k l), \tag{93}$$

provided that the operation (3) is applied on (92) in a formal way. Straightforward examples show that the latter proviso is not always legitimate.

**42.** The definitions of §41 will be used in §50 and §59–§60. For the present, only the notations of §40 will be needed.

Since the Lambertian formulation, (59), of the definition, (15), of $f'(n)$ is an identity in the positive parameter $r$ (in case of convergence), it can be replaced by the corresponding identity in a complex variable, $z$; that is, by

$$\sum f(n)z^n = F(z), \quad \text{where} \quad F(z) = \sum f'(n)z^n/(1 - z^n). \tag{94}$$

Correspondingly, the results derived from (59) for $M(f)$ can be transcribed to the case of an arbitrary $\lambda$ in (89). The simplest fact in this direction may be formulated as follows:

(*) If $f(n)$ has an amplitude function, and if $\alpha = \alpha(\lambda)$ is an abbreviation for the function $M(f_{-\lambda})$, where $0 \leqq \lambda < 1$, then (94) is convergent for $|z| < 1$, and $\alpha(\lambda)$ has the property that

$$(1 - r)F(z) \to \alpha(\lambda) \quad \text{as} \quad z \to e^{-2\pi i\lambda}, \qquad (|z| = r < 1), \tag{95}$$

where $\arg z = -2\pi\lambda$ is fixed or, more generally, $z$ approaches the circle $|z| = 1$ within a Stolz wedge. If (94) is convergent for $|z| < 1$, and if $f(n) = O_L(1)$ as $n \to \infty$, then $f(n)$ possesses an amplitude function, $M(f_\lambda) = \alpha(-\lambda)$, whenever

there exists on the range $0 \leqq \lambda < 1$ a function $\alpha = \alpha(\lambda)$ satisfying (95) for every fixed $\lambda$ in case of radial approach.

For every fixed $\lambda$, the first part of the assertion (*) is only a restatement of the fact that the summability $(C, 1)$ of a series suffices for the existence of the Abelian, and even of the Stolzian, limit of the power series generated by the series (the series to which this fact is to be applied is the one having (89) as its $n$-th partial sum). Correspondingly, the second part of the assertion (*) follows, just as in the proof of $(IX_2)$, from the elementary Tauberian theorem of Hardy and Littlewood (their theorem is to be applied to the two series having the real part of $f(n) + f_\lambda(n)$ and the imaginary part of $if(n) + f_\lambda(n)$ as their $n$-th partial sums, if $f \geqq 0$; and $f(n) = O_L(1)$ is, of course, reducible to $f \geqq 0$).

**43.** The first (that is, the purely Abelian) part of (*) leads to a short proof of an essential refinement of the so-called Knopp-Landau theorem in the theory of Lambert series.

In order to see this, let the coefficients of the Lambert series (94) be restricted by the condition $\Sigma \mid f'(n) \mid /n < \infty$. After it has repeatedly been observed that, if $\lambda$ is rational, say $\lambda = l/m$, where $1 \leqq l \leqq m$ and $(l, m) = 1$, then the limit $\alpha(\lambda)$ occurring in (95) must exist, and possess the value (82) belonging to the index $m$ thus defined, Landau [59] has shown by a somewhat lengthy (since Diophantine) proof that (95) is satisfied by $\alpha(\lambda) = 0$, if $\lambda$ is irrational; cf. Späth [71]. However, it is clear from $(80_3)$ and from the Abelian character of the proof of the trivial part of (*) in §42, that a much sharper result is contained in that corollary of (XVI) according to which (81)-(82) holds if (81) is meant in the sense of §40. In other words, the full force of (XVI) is not used, since the almost periodicity $(B)$ of (81) is not needed, and what results is nevertheless a Tauberian refinement of the Knopp-Landau theorem, although no Tauberian element, but only the Abelian lemma of Frobenius, occurs in this deduction.

**44.** If a function $f(n)$ is almost periodic $(B)$, its Fourier series need not be expressible by means of the Ramanujan sums $(80_3)$ in the form (81). In fact, if $\{\lambda_k\}$ is any real sequence, and if the $k$-th coefficient in (90) is chosen to be $\beta_k$, where $\{\beta_k\}$ is any sequence satisfying $\Sigma \mid \beta_k \mid < \infty$, then the series (90) converges to a uniformly almost periodic function having the series (90) as its Fourier series. In the general case, every $\lambda_k$ is rational if, but not only if, (90) can be contracted so as to have the form (81). In other words, (81) and $(80_3)$ require that the $\phi(m)$ values of $k$, which in virtue of $\lambda_k = l/m$ belong to a common $m$, should correspond to Fourier constants which are independent of $l$, where $(l, m) = 1$ and $1 \leqq l \leqq m$. According to the proof of (XVI), this condition is satisfied if the sequence of the partial sums of the Eratosthenian series (75) of $f(n)$ tends to $f(n)$ in the mean $(B)$.

As an illustration of the last remark, it is instructive to verify the second of the ssertions of (XV) by showing that the Fourier series $(B)$ of the function, $f(n)$, defined in the proof of (XV) is not of the form (81).

First, if $g(n)$ is any periodic function, and if $r$ denotes its period, then trigonometric interpolation supplies the Fourier expansion

$$(96) \qquad g(n) = \sum_{j=1}^{r} \alpha_j \exp(2\pi inj/r); \qquad \alpha_j = r^{-1} \sum_{n=1}^{r} g(n) \exp(-2\pi ijn/r).$$

Next, if $g(n)$ is 0 for $r - 1$ of the $r$ values of $n$ contained in a period, and if $g(s) = 1$, where $s$ denotes the exceptional $n$ between $n = 1$ and $n = r$, then (96) shows that $\alpha_j = r^{-1} \exp(-2\pi ijs/r)$, hence

$$(97) \qquad g(n) = r^{-1} \sum_{j=1}^{r} \exp(2\pi i[n - s]j/r).$$

It now suffices to apply (97) to each of the functions $g = g_1, g_2, \cdots$ defined in §29, and to observe that, in virtue of (79), the Fourier series $(B)$ of $f(n)$ can be obtained by formal addition of the Fourier series, (96), of $g = g_1, g_2, \cdots$.

**45.** As another illustration, suppose that (90) holds for a sequence of rational exponents $\lambda_k$, and replace, for every fixed $m$, the $\phi(m)$ simple vibrations $\exp(2\pi inl/m)$, where $(l, m) = 1$ and $1 \leqq l \leqq m$, by their $\phi(m)$ superpositions

$$(98) \qquad g_m^l(n) = \sum_{h=1}^{m} \exp(2\pi i[h^2 l + hn]/m),$$

[occurring in Cauchy's proof ([8]; cf. Lerch [60]) for the reciprocity law of the Gaussian sums (that is, of the sums (98) belonging to $n = m$), and differing from the sums $S_{pq\nu}$ of Kloosterman [55] only in notation].

If (98) is introduced into (90), where $l = l_k$, $m = m_k$ and $\lambda_k = l/m$ by assumption, then the Fourier series appears in the form

$$(99) \qquad f(n) \sim \sum_{n=1}^{\infty} \sum_{l}{}' \alpha_m^l g_m^l(n),$$

where the coefficients $\alpha$ represent the Fourier constants and the accent of the inner summation sign indicates the same $l$-range as in $(80_3)$. [Corresponding to the remark following (98), the Hardy-Littlewood solution of Waring's problem depends on Fourier series of a type generalizing (99), the polynomial in $h$ beneath the exponential sign in (98) being of degree $r$ in the problem of $r$-th powers. Incidentally, it is a trivial consequence of the results of Hardy and Littlewood that their trigonometric series obtained actually are Fourier series.]

In the case (99), it is seen from (98) and $(80_3)$ that what parallels Fourier series of the Ramanujan type (81) are those series in which each but the first of the $\phi(m)$ Fourier coefficients $\alpha_m^l$ belonging to a fixed $m$ happens to be 0 for every $m$. In fact, (99), (98) are then reduced to

$$(100) \quad f(n) \sim \sum_{m=1}^{\infty} \alpha_m g_m(n), \quad \text{where} \quad g_m(n) = \sum_{h=1}^{m} \exp(2\pi i[h^2 + hn]/m).$$

However, (100) involves for $f(n)$ restrictions essentially different from those imposed on $f(n)$ by (81), $(80_3)$. Actually, the formal assumption of (100)

replaces the *Eratosthenian approximations* (76) leading to (81), ($80_3$) by assumptions substantially expressible in terms of the *modul group* (cf. Petersson [66]). I expect to return to the resulting questions of almost periodic behavior.

## Multiplicative Functions

**46.** Let $f(n)$ be a multiplicative function. This means, by (39), that $f'(n)$ is multiplicative. Hence, $|f'(n)|/n$ is a multiplicative function, and so (38), (40) and (51) imply that $\Sigma\, |f'(n)|/n < \infty$ if and only if

$$(101) \qquad \sum_p \sum_{k=1}^{\infty} |f(p^k) - f(p^{k-1})|/p^k < \infty; \qquad f(1) = 1.$$

Since the number, $d(n)$, of the divisors of $n$ is the multiplicative function for which $d(p^k) = k + 1$, it is similarly seen that $\Sigma\, d(n)\, |f'(n)|/n < \infty$ if

$$(102) \qquad \sum_p \sum_{k=1}^{\infty} k\, |f(p^k) - f(p^{k-1})|/p^k < \infty; \qquad f(1) = 1.$$

Finally, $\Sigma\, |f'(n)| < \infty$ is equivalent to

$$(103) \qquad \sum_p \sum_{k=1}^{\infty} |f(p^k) - f(p^{k-1})| < \infty; \qquad f(1) = 1.$$

Consequently, (XVI), (XVIII) and the criterion (88) of §37 respectively imply that

(i) if (101) is satisfied, then $f(n)$ is almost periodic ($B$) and

$$(104) \qquad f(n) \sim \sum_{m=1}^{\infty} a_m c_m(n), \qquad a_m = \sum_{m \mid n} f'(n)/n,$$

where the series representing $a_1 = M(f)$, and therefore the series representing any of the Fourier constants, $a_m$, is absolutely convergent in virtue of (101);

(ii) if (101) is replaced by the more stringent assumption (102), then the Fourier series (104) is absolutely convergent, and has the sum $f(n)$, for every $n$;

(iii) if the drastic condition (103) is satisfied, then the Fourier series (104) is absolutely-uniformly convergent and has the sum $f(n)$, which implies in particular the uniform almost periodicity of the function $f(n)$.

**47.** As an illustration, consider the sum, $\sigma_\alpha(n)$, of the $\alpha$-th powers of all divisors, $d$, of $n$, where $d^\alpha$ is defined in terms of the real logarithm of $d$. Thus, if $\alpha$ is arbitrarily fixed,

$$(105) \qquad f(p^k) = \sum_{j=0}^{k} p^{-\alpha j} \quad \text{if} \quad f(n) = \sigma_\alpha(n)/n^\alpha.$$

Since $\sigma_\alpha(n)$, hence $f(n)$, is multiplicative, it follows from (ii) that $\sigma_\alpha(n)/n^\alpha$ is almost periodic ($B$), and is represented by its Fourier series, whenever the sums (105) satisfy (102), that is, whenever $\beta > 0$, where $\beta$ denotes the real part of $\alpha$. Since (37), (39) and (105) imply that

$$\sum_{m \mid n} f'(n)/n = \sum_{m \mid n} n^{-\alpha}/n = \sum_{m \mid n} n^{-\alpha-1} = m^{-\alpha-1}\zeta(\alpha + 1),$$

it follows from (104) that

$$\sigma_\alpha(n)/n^\alpha = \zeta(\alpha+1)\sum_{m=1}^{\infty} m^{-\alpha-1} c_m(n) \quad \text{if} \quad \beta > 0, \tag{106}$$

and that the trigonometric series (106) is the Fourier series $(B)$ of the function which it represents.

The identity (106) was discovered by Ramanujan ([69], p. 185) who used, of course, a different approach but made precisely the above assumption, $\beta > 0$. If $\beta > 1$, then, since $(80_3)$ implies that $|c_m(n)| \leqq m$, it is implied by Ramanujan's identity (106) that $\sigma_\alpha(n)/n^\alpha$ is uniformly almost periodic and has the series (106) as its Fourier series (incidentally, this, but nothing more than this, follows from (iii), §46). But the trigonometric series (106) turns out to be the Fourier series $(B)$ of the function represented by it also when $0 < \beta \leqq 1$; a fact which, in view of §38, is by no means implied by the truth of (106). That $f(n)$ cannot be uniformly almost periodic, and therefore the series (106) cannot be uniformly convergent, in the case $0 < \beta \leqq 1$, is clear from the fact that, the function defined by (105) being multiplicative, not even $f(n) = O(1)$ is true when $0 < \beta \leqq 1$.

**48.** The situation is similar if $\sigma_\alpha(n)$ in (105) is replaced by $\phi_\alpha(n)$, where $\phi_\alpha(n)$ is, for a fixed positive interger $\alpha$, Jordan's generalization of Euler's $\phi(n) = \phi_1(n)$; that is, $\phi_\alpha(n)$ denotes the number of $\alpha$-tuples of (not necessarily distinct) positive integers not exceeding $n$ and having a greatest common divisor relatively prime to $n$. Then, for every fixed $\alpha$,

$$f(p^k) = 1 - p^{-\alpha} \quad \text{if} \quad f(n) = \phi_\alpha(n)/n^\alpha, \tag{107}$$

and $f(n)$ is multiplicative. This extends the definition of $\phi_\alpha(n)$ to the case $\alpha \neq 1, 2, \cdots$. Let $\beta$ denote the real part of $\alpha$, and suppose that $\beta > 0$. Then, if (ii), §46 is applied as in §47, there results not only Ramanujan's identity ([69], p. 187),

$$\zeta(\alpha+1)\phi_\alpha(n)/n^\alpha = \sum_{m=1}^{\infty} a_m c_m(n); \qquad a_m = \mu(m)/\phi_{\alpha+1}(m), \quad (\beta > 0), \tag{108}$$

which parallels (106), but also the fact that the trigonometric series (108) is the Fourier series $(B)$ of the function which it represents. As in §47, this follows from Ramanujan's identity if and only if $\beta > 0$ is replaced by $\beta > 1$. But the assumption $\beta > 1$ makes the situation trivial and excludes Euler's $\phi(n)$, since the latter belongs to $\alpha = 1$.

**49.** Actually, $\phi_\alpha(n)/n$ cannot be uniformly almost periodic, hence the series (108) cannot be uniformly convergent, in Euler's case.

Since $\phi(n)/n = O(1)$, this negation cannot be proved by the argument applied at the end of §47. However, it is clear from the definition of $G(f)$ in §3, that the uniform almost periodicity of $f(n)$ is sufficient for the existence of the Gaussian density, $G(f)$. But it is easily verified from elementary properties of the distri-

bution of primes, that the Gaussian density cannot exist for the function $\phi(n)/n = \Pi(1 - p^{-1})$, where $p \mid n$.

Accordingly, the necessary conditions (i)–(ii), §4 for the existence of $G(f)$ are satisfied by the function $f(n) = \phi(n)/n$, but $G(f)$ does not exist, that is, a trivial necessary condition of uniform almost periodicity is violated. In fact, the non-existence of $G(f)$ implies that $f(n)$ cannot even be almost periodic ($W$). In this connection, cf. Hartman and Wintner [43].

If $\psi(n)$ denotes, as usual, the degree of the modular equation of the $n$-th *Stufe* (that is, the number of the transformation subgroups of order $n$ in the modul group), then, according to Dedekind, $\phi(n)/n = \Pi(1 - p^{-1})$ must be replaced by $\psi(n)/n = \Pi(1 + p^{-1})$, where $p \mid n$. Hence it is clear that the situation is precisely the same as in the case $\alpha = 1$ of (105), and is therefore apparently less favorable than, but actually about the same as, in the case $\alpha = 1$ of (107).

**50.** A case of uniform almost periodicity presents itself in what corresponds to coefficients of auto-correlation (cf. §41) of Euler's $\phi(n)$ itself (instead of $\phi(n)/n$). In fact, if $l$ is a fixed positive number, then, according to Ingham [45],

$$(109) \qquad \sum_{m=1}^{n} \phi(m)\phi(m + l) \sim cf(l)n^3 \quad \text{and} \quad \sum_{m=1}^{n-1} \phi(m)\phi(n - m) \sim \tfrac{1}{2}cf(n)n^3$$

as $n \to \infty$, where the function $f(n)$ and the constant $c$ are given by

$$(110) \qquad f(n) = \prod_{p \mid n} (p^3 - 2p + 1)/(p^3 - 2p) \quad \text{and} \quad 3c = \prod_{p} (1 - 2p^{-2})$$

respectively. But it is clear from (110) that $f(n)$ is the multiplicative function for which $f(p^k)$ is $1 + (p^3 - 2p)^{-1}$. Hence the assumption, (103), of (iii), §46 is satisfied.

This example is of interest, since, as a rule, functions $f(n)$ that are uniformly almost periodic are either too artificial from an arithmetic point of view or uniformly almost periodic practically by definition. In other words, the condition of uniform almost periodicity happens to be too restrictive to have arithmetical applications of a non-trivial nature.

**51.** The genesis of this situation can well be illustrated by a function considered by Gram [29] which, though instructive in itself, seems to have received just as little attention as a similar construction of Dedekind [13].

Gram's purpose appears to have been (cf. his subsequent papers [30], [31]) an elementary model of the mysterious mechanism which leads to Riemann's explicit formula in the distribution theory of primes. Incidentally, Dedekind's parallel construction serves the same purpose (even though there is no evidence that Dedekind was aware of this; cf. Wintner [95]). Gram proves that, if $F(n) = n - F([\frac{1}{2}n])$, then $F(n)/n \to \frac{2}{3}$ as $n \to \infty$, and his formulae lead him to suspect, and perhaps even to claim, that the remainder term $F(n) - 2n/3$ has a principal "wobble" (cf. Hardy [36], Chap. II). The following Fourier analysis

will imply that the resulting approximation is almost as misleading as, according to Littlewood, the principal wobble of Riemann's series is.

Corresponding to the Euclidean algorithm of dyadic fractions, successive application of the definition, $F(n) = n - F([\frac{1}{2}n])$, gives

$$F(n) = \sum_{j=0}^{\infty} (-1)^j [2^{-j} n] = \sum_{j=0}^{[\log n]} (-1)^j [2^{-j} n], \tag{111}$$

since $[2^{-j}n] = 0$ for every $j > [\log n]$, where the logarithm is of base 2. A comparison of (111) with (19) gives

$$F(n) = \sum_{m=1}^{n} f(m), \tag{112}$$

if $f(n)$ denotes the function for which $f'(n)$ is $(-1)^j$ or 0 according as $n$ is or is not of the form $2^j$, where $j = 0, 1 \cdots$. In other words, $f'(n)$ is the multiplicative function for which

$$f'(p^k) = 0 \quad \text{if} \quad p \neq 2 \quad \text{and} \quad f'(2^k) = (-1)^k. \tag{113}$$

Hence (39) and (40) show that $f(n)$ is multiplicative and

$$f(p^k) = 1 \quad \text{if} \quad p \neq 2 \quad \text{and} \quad f(2^{2k-1}) = 0, \qquad f(2^{2k}) = 1. \tag{114}$$

This means that $f(n)$ is 0 or 1 according as the exponent of the highest power of 2 dividing $n$ is odd or even (it is even if it is 0). In particular, the multiplicative function $f(n)$ is the characteristic function of a set $S$ (cf. the beginning of §29). All of this is also clear from the dyadic interpretation, alluded to before (111); an interpretation which seems to have escaped Gram [29].

According to (111) and (112),

$$\sum_{m=1}^{n} f(m) = n \sum_{j=0}^{[\log n]} (-2^{-1})^j + O(\log n) = n \sum_{j=0}^{\infty} (-2^{-1})^j + O(\log n), \tag{115}$$

which, by (3), implies that $M(f)$ exists, its value being the sum of the last series, that is, $\frac{2}{3}$. Furthermore, since (114) satisfies the assumption, (102), of (ii), §46, the function $f(n)$ is almost periodic $(B)$ and is represented by its Fourier series; the latter results in an explicit form, if the values attained by the multiplicative function $f'(n)$ are substituted from (113) into (104). Actually, as it will turn out in §52, the almost periodicity $(B)$ of $f(n)$ has nothing to do with the explicit structure of the set $S$ having the multiplicative function $f(n)$ as its characteristic function.

It is clear from (113) that the multiplicative function $f(n)$ is 2-adic, if a function $f(n)$ is called $p$-adic when $f'(n)$ vanishes for every $n$ distinct from $1, p, p^2, \cdots$ (and possibly for some of the integers $1, p, p^2, \cdots$ also), where $p$ is a fixed prime. According to Toeplitz [79], a $p$-adic function of $n$ is uniformly almost periodic if and only if its value attained at $n = p^k$ tends to a limit as $k \to \infty$. On the other hand, a $p$-adic function of $n$ is almost periodic $(W)$ if and only if it is bounded (Hartman and Wintner [43]). Thus it is clear from (114) that $f(n)$ cannot be uniformly almost periodic and that it is almost periodic $(W)$.

**52.** Let a set, $S$, of positive integers be called multiplicative if its characteristic function, $f(n)$, is a multiplicative function. Clearly, a multiplicative set $S$ is given if and only if it is known for every prime power, $p^k$, where $k = 1, 2, \cdots$, whether $n = p^k$ is or is not in $S$; the value of $f(p^k)$ being 1 in the first case and 0 in the second case. It is understood that the decision of this alternative is an arbitrary function of the two variables $p$, $k$, and that $n = 1$ is contained in every multiplicative $S$.

Instances of finite multiplicative sets are the set of the square-free divisors of $m$ and the set of all divisors of $m$, where $m$ is arbitrarily fixed. The characteristic function of the latter set is $e_n(m)/n$ in terms of the notation $(80_1)$. The infinite multiplicative set consisting of those positive integers $n$ which are relatively prime to a fixed $m$ has the characteristic function $\rho(m, n)$, where $\rho(m, n) = \rho(n, m)$ is the function considered in Kronecker's *Vorlesungen*, pp. 246–250.

Clearly, the assumption of (iii), §46 is satisfied by the characteristic function of a multiplicative set $S$ only if the double series (103) has a finite number of non-vanishing terms (which implies that all but a finite number of primes must be in $S$). It is easy to see that this trivial sufficient condition for the uniform almost periodicity of $f(n)$ is necessary as well (the reasons are similar to those which preclude for $\phi(n)/n$ the existence of a Gaussian density). In view of the example of §51, the conditions for the almost periodicity $(W)$ of $f(n)$ are clearly more favorable, although $f(n)$ cannot be almost periodic $(W)$ for an arbitrary multiplicative $S$. However, it is always almost periodic $(B)$.

In order to prove this, let a prime, $p$, be called a $q$ or an $r$ according as it is or is not contained in $S$. Two cases will be distinguished, according as $\Sigma\, r^{-1} < \infty$ or $\Sigma\, r^{-1} = \infty$. In the first case, the assumption of (ii), §46 is satisfied, since (102) is always majorized by

$$\sum_q |1 - 1|q^{-1} + \sum_r |0 - 1|r^{-1} + \sum_p \sum_{k=2}^{\infty} k(1 + 1)p^{-k} = \sum_r r^{-1} + \text{const.}$$

Hence, in the first case $f(n)$ not only is almost periodic $(B)$ but it also has a Fourier series satisfying (86) for every $n$. This cannot be true in the second case, since if $S$ consists of $n = 1$ alone, then $f(n)$ is certainly almost periodic $(B)$, and has the Fourier series $0 + 0 + \cdots$, although $f(1) = 1$; cf. (iii), §38. Actually, this is always true in the second case. In fact, it will be verified in §79 that the mean $M(f)$ exists and is 0 whenever $\Sigma\, r^{-1} = \infty$. But $M(f) = 0$ and $f(n) \geqq 0$ imply that $f(n)$ is almost periodic $(B)$, with $0 + 0 + \cdots$ as Fourier series.

**53.** An extreme instance of the first case is the characteristic function, $|\mu(n)|$, of the multiplicative set consisting of all square-free positive integers. Since $|\mu(p)| = 1$ for every $p$ and $|\mu(p^k)| = 0$ whenever $k > 1$, it follows from (17) and (18) that $|\mu(p^k)|' = 0$ for every $k \neq 2$ and $|\mu(p^2)|' = -1$. Hence

$$(116) \qquad |\mu(n)|' = \mu(n^{\frac{1}{2}}), \quad \text{where} \quad \mu(x) = 0 \quad \text{for} \quad x \neq [x],$$

as in (23). Substitution of (116) into (104) gives the Fourier series $(B)$ of the characteristic function $|\mu(n)|$.

An extreme instance of the second case is the following dual of the preceding example: Let $S_0$ be the multiplicative set consisting of those positive integers, say of $1 = h_1 < h_2 < \cdots$, which have *no simple* prime factor. Thus if $\theta_0(n)$ denotes the characteristic function of $S_0$, then $\theta_0(p) = 0$ and $\theta_0(p^2) = 1$, $\theta_0(p^3) = 1, \cdots$. Hence, by Euler's factorization,

$$\begin{aligned}(117)\qquad \sum h_n^{-s} &\equiv \sum \theta_0(n)n^{-s}\\ &= \prod (1 + 0 + p^{-2s} + p^{-3s} + \cdots) = \zeta(2s)\zeta(3s)/\zeta(6s),\end{aligned}$$

if $\sigma > \frac{1}{2}$ (cf. Wintner [94]). In fact, the value of $1 + p^{-2s} + p^{-3s} + \cdots$, being equal to $1 + p^{-2s}/(1 - p^{-s})$, is identical with the reciprocal value of

$$(1 - p^{-2s})(1 - p^{-3s})/(1 - p^{-6s}).$$

**54.** For the Dirichlet generator of the multiplicative function $\gamma(n)$ defined by $\gamma(p^k) = k$, Cantor [6] observed an identity of the same type as (117), namely

$$(118)\qquad \sum \gamma(n)n^{-s} = \zeta(s)\zeta(2s)\zeta(3s)/\zeta(6s),$$

where $\sigma > 1$. He did not point out that $\gamma(n)$ has a simple significance, which can be realized as follows:

If $p, q, \cdots$ are the distinct prime divisors of an integer $n = p^k q^l \cdots$, where $k > 0, l > 0, \cdots$ and $n \neq 1$, let $(\alpha, \beta, \cdots)$, $(\lambda, \mu, \cdots)$, $\cdots$ denote any set of sets of positive integers satisfying $\alpha + \beta + \cdots = k, \lambda + \mu + \cdots = l, \cdots$ and $\alpha \leqq \beta \leqq \cdots, \lambda \leqq \mu \leqq \cdots, \cdots$. Then the description of all finite Abelian groups in terms of primary bases asserts that every Abelian group of order $n$ is simply isomorphic to the group of ordinary multiplication in exactly one collection $\{1; p^\alpha, p^\beta, \cdots; q^\lambda, q^\mu, \cdots; \cdots\}$, and vice versa. But the number of all these collections belonging to a fixed $n = p^k q^l \cdots$ is $(\alpha + \beta + \cdots)\cdot(\lambda + \mu + \cdots)\cdot \cdots$, that is, $kl \cdots$. Hence, the number of all Abelian groups of order $n$ is $kl \cdots$, if $n \neq 1$. Since it is 1 if $n = 1$, the multiplicative function $\gamma(n)$ defined before (118) represents the number of the Abelian groups of order $n$.

It follows that the number of all Abelian groups of order $n$, when considered as a function of $n$, is almost periodic $(B)$ and is represented by its Fourier series for every $n$. In fact, it is clear from $f(p^0) = f(1) = 1$ that the assumption, (102), of (ii), §46 is satisfied if $f(p^k) = k$. On the other hand, the necessary condition (i), §4 for the existence of $G(f)$ is not satisfied by $f = \gamma$.

It is seen from (117), (118) and (25) that $\gamma'(n) = \theta_0(n)$. Hence, (104) shows that the Fourier constant, $a_m$, of $\gamma(n)$ has the value

$$(119)\qquad a_m = \sum_{m \mid n} \theta_0(n)/n = \sum_{m \mid h_n} h_n^{-1},$$

by (117), where $\sigma > \frac{1}{2}$. In particular, since $a_1 = \Sigma f'(n)/n$ is the mean, $M(f)$, of $f = \gamma$, it is clear from (119) and (117) that, in accordance with Cantor's identity (118),

$$M(\gamma) = \zeta(2)\zeta(3)/\zeta(6). \tag{120}$$

Since $\gamma(n)$ is at least 1 for every $n$ and becomes the larger the more composite $n$ is, it might first appear paradoxical that the asymptotic average value of the number of the finite Abelian groups be so small a value as (120); a value which is barely 2, since $\zeta(2) = \pi^2/6$, $\zeta(6) = \pi^6/945$ and, according to Gram's table ([30], p. 269 or [32]), $\zeta(3) = 1.2020569 \cdots$. The paradox is cleared up by observing that $\gamma(n) = 1$ if and only if $n = i$, where $i$ denotes a square-free number; that $\gamma(n) = 2$ if and only if $n = p^2 i$, where $(p, i) = 1$; and so on.

## Part III

## THE STATISTICS OF THE PRIME NUMBER THEOREM



### The Distribution of Primes and Fourier Constants

**55.** In view of the examples (i)-(ii), §38, there arises the question as to the existence of a Fourier expansion for the function, $f(n) = \Lambda(n)$, on which the standard proofs of the prime number theorem depend; that is, for the function which, according to (26) and (15), can be defined by

$$\Lambda(n) = \delta_{np^k} \log p, \tag{121}$$

if $\delta_{nm} = 0$ when $n \neq m$ and $\delta_{nn} = 1$. It is indicated by a comparison of §41 with the approach of Hardy-Littlewood [38], [39] to Goldbach's problem (cf. §59-§61 below), that the Fourier expansion (90), if any, will be of the particular type of Ramanujan series (cf. §44-§45), that is, of the form

$$\Lambda(n) \sim \sum_{m=1}^{\infty} a_m c_m(n). \tag{122}$$

According to ($80_3$) and to the pair of conditions which defined (90) in §40, the assumption (122) means that

(i) if $\lambda$ is rational, say $\lambda = l/m$, where $(l, m) = 1$ and $1 \leqq l \leqq m$, then the Fourier constant $M(\Lambda_\lambda)$, where

$$\Lambda_\lambda(n) = \Lambda(n) \exp(-2\pi i \lambda n), \tag{123}$$

exists and has a value, $a_m$, that is independent of the $\phi(m)$ possible choices of $l$;

(ii) if $\lambda$ is irrational, the Fourier constant $M(\Lambda_\lambda)$ of (121) exists and vanishes.

It will be shown that, in this sense, (122) can *almost* be proved, without recourse to any form of Riemann's hypothesis. The italicized reservation refers to the possible violation of (ii) by a zero set of irrationals. It will remain undecided whether or not this $\lambda$-set of measure 0 is actually vacuous (in this connection, cf. Vinogradow [83], Theorem 1).

(*) *The prime number theorem, when formulated for all arithmetic progressions, implies* (*and is substantially equivalent to*) *the fact that the Fourier analysis* (122) *holds in the sense of* (i) *alone, and that the Fourier constants have the values*

$$a_m = \mu(m)/\phi(m), \qquad (m = 1, 2, \cdots). \tag{124}$$

*Furthermore,* (ii) *is true at any rate for almost all* $\lambda$.

According to ($80_3$), all frequencies $\lambda_k$ occurring in (90) when the series (90) is of the Ramanujan type (122), have rational values, and are therefore taken

care of by (i) alone. Correspondingly, the rôle of (ii) is that of assuring the completeness of the Fourier analysis. Thus the first part of (*) is interesting in itself.

The zero set of exceptional irrationals $\lambda$, if any, cannot have anything to do with the fact that, as observed by Hardy [35] in connection with Ramanujan's result quoted in (ii), §38, the series obtained by substituting (124) into (122) converges to $\phi(n)/n$ times $\Lambda(n)$, and not to $\Lambda(n)$ itself, if* $n \neq 1$. In fact, this effect of discontinuity (cf. §16) depends only on the illegitimacy of the rearrangements mentioned at the beginning of §38.

**56.** In order to prove the first part of the italicized statement of §55, let

$$(126) \qquad \epsilon = \exp(-2\pi i\lambda), \quad \text{where } \lambda = l/m; \quad (l, m) = 1, \qquad 1 \leqq l \leqq m,$$

$l$ and $m$ being fixed. Thus, from (121),

$$(127) \qquad \sum_{n<x} \epsilon^n \Lambda(n) = \sum_{p<x} \epsilon^p \log p + \sum_{p^2<x} \epsilon^{p^2} \log p + \cdots.$$

According to (3), (123) and (126), the assertion (i), §55 is that the expression (127) is of the form $a_m x + o(x)$ as $x \to \infty$, where $a_m$ is supposed to have the value (124). But the number of the non-vacuous sums on the right of (127) is $O(\log x)$, and it is clear from $|\epsilon| = 1$ that, in view of Chebyshev's estimate $\Lambda(1) + \cdots + \Lambda([x]) = O(x)$, the $k$-th of these sums is majorized by a fixed multiple of $x^{1/k}$. Thus the total contribution of all sums but the first is $O(\log x)O(x^{\frac{1}{2}}) = o(x)$. Hence it is sufficient to show that the first sum on the right of (127) is $a_m x + o(x)$.

To this end, let $P^0$ denote the set of the prime divisors of $m$ (for the value of $m$ fixed above), and let all primes not contained in $P^0$ be classified into $\phi(m)$ subsequences $P_1, P_2, \cdots$ in such a way that $p$ is in $P_j$ if and only if $p \equiv l_j \pmod{m}$, where $l_1, l_2, \cdots$ represent the $\phi(m)$ residue classes relatively prime to $m$. Then, since (126) is a primitive $m$-th root of unity,

$$(128) \qquad \sum_{p<x} \epsilon^p \log p = O(1) + \sum_{j=1}^{\phi(m)} \epsilon^{l_j} \vartheta_j(x),$$

where $O(1)$ represents the contribution of the finite class $P^0$ and $\vartheta_j(x)$ denotes the sum of the logarithms of those primes contained in $P_j$ which are less than $x$. Thus the prime number theorem of the arithmetic progressions asserts that

* In terms of the above-mentioned parallelism between series $\Sigma\, a_m c_m(n)$ and periodic trigonometric series on the interval $0 \leqq t < 2\pi$, the proviso $n \neq 1$ can be compared to the proviso $t \neq 0$ in $f(t) = \Sigma\, \alpha_m \cos mt$ if $f(0) = 0$ and, for instance, $\alpha_{m-1} = 1/\log m$ (or, for that matter, $\alpha_m = 1/m$). In fact, it is seen from ($80_6$) and (124) that the series (122) is reduced at $n = 1$ to

$$(125) \qquad \Sigma\, c_m(1)\mu(m)/\phi(m) = \Sigma\, \mu(m)^2/\phi(m) \geqq \Sigma\, 1/\phi(p) = \infty,$$

although $\Lambda(1) = 0$, by (121).

$\vartheta_j(x) \sim x/\phi(m)$ as $x \to \infty$ holds for each of the $\phi(m)$ values of $j$. On the other hand, $\epsilon^{l_1}, \epsilon^{l_2}, \cdots$ are the $\phi(m)$ primitive $m$-th roots of unity, and so their sum is $\mu(m)$. It follows therefore from (128) that the first sum on the right of (127) is $a_m x + o(x)$, if $a_m$ is defined by (124).

**57.** The proof of (i) thus completed follows the pattern of an elementary result of Fatou [25] for which, instead of the prime number theorem, only Mertens' elementary formula is needed.

If log $p$ on the right of (127) is replaced by $1/p$, then, according to Mertens, each of the $\phi(m)$ functions $\vartheta_j(x)$ on the right of (128) must be replaced by $(\log\log x)/\phi(m) + O(1)$, where each of the $\phi(m)$ functions $O(1)$ of $x$ tends to a limit as $x \to \infty$. On the other hand, the term corresponding to the $O(1)$ in (128) is independent of $x$ for large $x$, since it represents the contribution of a finite class, $P^0$, of primes. It follows therefore from the absolute convergence of the double series

$$\sum_p \sum_{k=2}^{\infty} \epsilon^{p^k}/p^k, \qquad \left(\sum_p \sum_{k=2}^{\infty} 1/p^k < \infty\right),$$

that, if $\epsilon$ and $a_m$ are defined by (126) and (124),

$$\sum_{p<x} \epsilon^p/p = a_m \log\log x + \text{Const.} + o(1) \quad \text{as} \quad x \to \infty.$$

Since the value (124) is 0 only when $m$ is not square-free, it follows that the power series $\Sigma\, z^p/p$ is convergent or divergent at the rational point $z = \epsilon$ of the circle $|z| = 1$ according as the integer $m$ defined by (126) is not or is square-free. This is the elementary result of Fatou.

It is also seen that, if $m$ is square-free, then there exists a real constant such that $\Sigma\, z^p/p$ diverges either to $+\infty + i$ const. or to $-\infty + i$ const. at the rational point $z = \epsilon$, the sign of $\infty$ being the same as that of $\mu(m)$. But the rational points $z = \epsilon$ satisfying (126) and either of the restrictions $\mu(m) = \pm 1$ lie dense on the circle $|z| = 1$. It follows therefore from a standard theorem on series of continuous functions, that those irrational points on $|z| = 1$ at which the power series $\Sigma\, z^p/p$ is divergent, with a divergence of the type $+\infty + ic$, form a set of the second category (and so, in particular, a set that is non-enumerable on every subinterval), and that the same is true if $+\infty + ic$ is replaced by $-\infty + ic$.

However, the power series is convergent almost everywhere on $|z| = 1$. This follows for $\Sigma\, z^p/p$ in the same way as for the weirder function $\Sigma\, z^p(\log p)/p$ in §58 below.

It would be interesting to know whether or not the Fourier series $\Sigma\, (\sin pt)/p$, which obviously is of class $(L^2)$ and represents the imaginary part of $\Sigma\, z^p/p$ on $|z| = 1$, is convergent for every $t$ (instead of almost every $t$).

**58.** The second part of the assertion (*), §55 has nothing to do with prime numbers, since it is implied by a general lemma on Fourier series of class $(L^2)$.

The lemma in question, when particularized to the case of power series, states

that the existence of a $\delta > 0$ satisfying $\Sigma \mid h(n) \mid^2 n^\delta < \infty$ (and, incidentally, even the condition $\Sigma \mid h(n) \mid^2 \log n < \infty$) is sufficient to guarantee the convergence of $\Sigma\, h(n)z^n$ almost everywhere on $|z| = 1$. Hence, if $h(n) = \Lambda(n)/n$ and $z = \exp(it)$, an application of (4) to the function $g(n) = nh(n)z^n$ of $n$ shows that, in order to prove the existence and the vanishing of the mean, $M(f_\lambda)$, of (89) for almost all values of $\lambda$ in the case $f(n) = \Lambda(n)$, it is sufficient to ascertain that $\Sigma \mid \Lambda(n)/n \mid^2 n^\delta < \infty$ holds for some $\delta > 0$. But is it clear from (121) that this condition is satisfied by every $\delta < 1$.

Incidentally, (121) implies that the sufficient conditions of W. H. Young for the common part of all Lebesgue classes $(L^2)$, $(L^3)$, $\cdots$ (and even the corresponding condition for an exponential class) are also satisfied.

**59.** Suppose for a moment that the zero set of $\lambda$-values, which are excluded in the second part of (*), §55, is vacuous (or at any rate irrelevant). Then the assertion of (*), §55 becomes that (122) and (124) represent for $f(n) = \Lambda(n)$ a Fourier expansion in the sense defined after (90). Suppose in addition that the formal transition from (90) to (93) is legitimate in this case (actually, this second hypothesis, which is independent of the first, is all that is relevant in the sequel). Then $(80_3)$ shows that, if $l$ in (93) is replaced by $n$,

$$M(\Lambda^{(n)}) = \sum_{m=1}^{\infty} |a_m|^2 c_m(n), \quad \text{where} \quad a_m = \mu(m)/\phi(m), \tag{129}$$

by (124). In the case $l = 0$ excluded after (92), it is easily seen from (121) that the mean, (3), of the function (93) belonging to $f = \Lambda$, $l = 0$ does not exist, since

$$M(\Lambda^2) = M(|\Lambda|^2) = M(\Lambda^{(0)}) = \infty \tag{130}$$

(on the other hand, the prime number theorem asserts that $M(\Lambda)$ exists; its value is 1).

For a fixed positive $n$, let $g(m)$ denote the $m$-th term of the series (129). Then $g(m)$ is multiplicative, since all three functions $\mu(m)$, $\phi(m)$, $c_m$ of $m$ are. Furthermore, since $\mu(p^k) = 0$ if $k > 1$, and $\mu(p) = -1$, it is clear that $g(p^k) = 0$ if $k > 1$, and that substitution of $(80_7)$, where $p - 1 = \phi(p)$, gives $g(p) = 1/\phi(p)$ or $g(p) = -1/\phi(p)^2$ according as $p$ does or does not divide the fixed integer $n$. It follows therefore from $\phi(p) = p - 1$ and from $\Sigma\, p^{-2} < \infty$, that the sufficient condition, (51), for the validity of (50) is satisfied. Consequently, (129) is equivalent to

$$M(\Lambda^{(n)}) = \prod_{p \mid n} \{1 + (p-1)^{-1}\} \prod_{p \nmid n} \{1 - (p-1)^{-2}\}.$$

Since $1 - (2 - 1)^{-2} = 0$, the first factor of the second product vanishes for every odd $n$. Hence it is easily verified that the last relation can be written in the form

$$M(\Lambda^{(2n)}) = 2Af(n), \qquad M(\Lambda^{(2n-1)}) = 0, \tag{131}$$

if $A$ is the positive constant

$$A = \prod_{p>2} \{1 - (p-1)^{-2}\} \tag{132}$$

and $f(n)$ denotes the function $\prod_{2<p|2n} \frac{p-1}{p-2}$. Thus

$$f(p^k) = (p-1)/(p-2) \quad \text{if} \quad p \neq 2 \quad \text{and} \quad f(2^k) = 1 \tag{133}$$

and

$$f(nm) = f(n)f(m) \quad \text{whenever} \quad (n, m) = 1; \qquad f(1) = 1. \tag{133 bis}$$

**60.** The multiplicative function $f(n)$ defined by (133) happens to be identical with Sylvester's oscillatory function in Goldbach's problem.

In fact, if $\rho(n)$ denotes the number of the representations of $n$ as a sum of two primes, $p$ and $q$, then the assertion of the rule of Sylvester [72], [75] is that

$$\rho(2n) \sim 2Af(n)n/\log^2 n \quad \text{as} \quad n \to \infty; \qquad \rho(2n-1) = 0. \tag{134}$$

For reasons explained by Hardy and Littlewood ([38], [39]), Sylvester suggested, instead of (132), an erroneous numerical value for the constant $A$ occurring in (134). The appearance of $n/\log^2 n$ in (134) is easy to guess, since, the sum $\rho(1) + \cdots + \rho(2n)$ being the number of pairs of primes satisfying $p + q \leqq 2n$, the formulation $\pi(x) \sim x/\log x$ of the prime number theorem implies that

$$\frac{1}{2n} \sum_{m=1}^{2n} \rho(m) = \frac{1}{2n} \sum_{p \leqq 2n} \pi(2n - p) \sim \frac{n}{\log^2 n} \quad \text{as} \quad n \to \infty, \tag{135}$$

since $p_m \sim m \log m$ as $m \to \infty$. What is ingenious in Sylvester's law, (134), is the discovery of the explicit form, (133)-(133 bis), of the delicate function $f(n)$, which exhibits the "irregularities" of $\rho(n)$. In view of the following theorem, it appears to have an historical interest that Sylvester states the function $f(n)$ to be expressible in terms of infinite series involving roots of unity of arbitrarily high degree, and that, as a contradistinction, he refers to the reduction factor, (135), as *non*-periodic.

(**) *The function* (129) *of* $n$, *representing the formal auto-correlation coefficient,* $M(\Lambda^{(n)})$, *of Chebyshev's function* (121), *is connected with Sylvester's law* (134) *by the relations* (131)-(134), *where the multiplicative function* $f(n)$, *defined by* (133), *is almost periodic* ($B$) *and has a Fourier series of the form* (81)-(82) *which, in addition, converges to* $f(n)$ *for every* $n$.

**61.** The situation in (**) is as follows. As shown by van der Corput [11], the method of Vinogradow [83] proves the truth of Goldbach's conjecture for all but $o(n)$ of the even integers not exceeding $n$ (previously, Hardy and Littlewood [40] proved this under a modified form of Riemann's hypothesis). However, Sylvester's law has never been proved, not even under the assumption of the truth of Riemann's hypothesis for Dirichlet's $L$-series. But Vinogradow's

method supplies the truth of the analogue of (134) not only for the case of 3 primes, treated loc. cit. [83], but for the case of $j$ primes as well, provided that $j > 2$ (previously, Hardy and Littlewood [39] proved the same under a modified form of Riemann's hypothesis).

Goldbach's case, $j = 2$, has its particular difficulties in view of a coalescence of "major" and "minor" arcs. However, *the formal "singular series" belonging to* $j = 2$ *proves to be identical with that Fourier series* $(B)$ *to which* (**) *refers*; a series which, by $(80_3)$, is of the type described by Sylvester and represents, in accordance with his contradistinction, a function that is almost periodic $(B)$. In view of §38, this is by no means obvious from the convergence of the trigonometric series representing $f(n)$. In fact, since $\Pi(1 + p^{-1}) = \infty$ and (133)–(133 bis) imply that $f(n) \neq O(1)$, it is obvious that the series (86) cannot now be uniformly convergent. Accordingly, the principal content of (**) is that the trigonometric series (86) now happens to be the Fourier series $(B)$ of the function which it represents.

Needless to say, not even the almost periodicity $(B)$ of $f(n)$ can be inferred from the convergence of the singular series. Actually, the almost periodicity $(B)$ of $f(n)$ shows that the standard description of $f(n)$ as the contribution of the "irregularities" of Goldbach's problem is rather misleading, since $f(n)$ admits of quite a regular harmonic analysis; in this connection, cf. Wintner [91].

In order to prove not only the almost periodicity $(B)$ of $f(n)$ but the additional assertion of (**), §60 as well, it is sufficient to observe that the assumption (102) of (ii), §46 is satisfied by (133). In fact, the series (102) is reduced to $\Sigma \mid f(p) - 1 \mid /p$, where $\mid f(p) - 1 \mid /p \sim 1/p^2$, by (133).

It may be mentioned that, in accordance with the end of §50, the problem of almost periodicity is trivial if Goldbach's problem is replaced by the problem of $j > 2$ primes, since then the sum of the singular series is uniformly almost periodic by virtue of its absolutely-uniform convergence.

The non-oscillatory factor, which is the generalization of (135) for $j > 2$, is just a reflection of the prime number theorem for every fixed $j$. In the next chapter, certain variants of the asymptotic laws of these non-oscillating factors will be considered.

## Poisson's Law and the Distribution of Primes

**62.** If $S$ is any set of positive integers, let $S(n)$ denote the number of those of its elements which do not exceed $n$. Thus, if $f_S(n)$ denotes the characteristic function of $S$,

$$S(n) = \sum_{m=1}^{n} f_S(m). \tag{136}$$

The set $S = S_0$ and its characteristic function $f_S(n) = \theta_0(n)$ considered in the second part of §53 represent the case $l = 0$ of the set $S = S_l$ and of its characteristic function $f_S(n) = \theta_l(n)$, if $S_l$ is defined as follows: A positive integer is in $S_l$ if and only if the number of its *simple* prime factors is exactly $l$. For instance,

$1, p, p^2, pq, p^2q$ are in $S_0$, $S_1$, $S_0$, $S_2$, $S_1$ respectively, if $q$ is a prime distinct from $p$. Since every $n$ is in exactly one of the sets $S_0$, $S_1$, $\cdots$, it is clear that

$$\sum_{l=0}^{\infty} \theta_l(n) = 1 \quad \text{for every } n. \tag{137$_1$}$$

On the other hand, from (136),

$$\sum_{m=1}^{n} \theta_l(m) = S_l(n) \quad \text{for every } n, \qquad (l = 0, 1, \cdots). \tag{137$_2$}$$

The bulk of the positive integers is certainly not in $S_0$. In fact,

$$S_0(n)/n = O(n^{-\frac{1}{2}+\epsilon}) \tag{139$_0$}$$

holds for every $\epsilon > 0$, since, the ordinary Dirichlet series (117) being convergent for $\sigma > \frac{1}{2}$, the sum function (137$_2$) belonging to $l = 0$ must be $O(n^{\frac{1}{2}+\epsilon})$ for every $\epsilon > 0$. On the other hand, if $A$ denotes the constant (120), that is,

$$A = \zeta(2)\zeta(3)/\zeta(6), \tag{138}$$

then

$$\frac{S_l(n)}{n} \sim \frac{A}{(l-1)!} \frac{(\log\log n)^{l-1}}{\log n}, \quad \text{if} \quad l \neq 0, \qquad (n \to \infty), \tag{139}$$

where $l$ is fixed (cf. Wintner [94]). According to (139$_0$) and (139), the relative frequencies of the integers contained in $S_0$ and in any fixed $S_l \neq S_0$ deviate almost by the order of $n^{-\frac{1}{2}}$. Incidentally, the estimate (139$_0$) can be refined to the asymptotic formula

$$S_0(n)/n \sim \text{const. } n^{-\frac{1}{2}}; \qquad \text{const.} = \zeta(\tfrac{3}{2})/\zeta(3), \tag{140}$$

which, in view of (138), may be imagined to be the limiting case of (139).

**63.** In order to verify these asymptotic laws, let $R_l$ denote the set of those square-free positive integers which are composed of exactly $l$ prime factors. For instance, $R_0$ consists of the single integer 1, and $R_1$ of all primes. It was apparently known already to Gauss [28] that the truth of the asymptotic formula

$$\frac{R_l(n)}{n} \sim \frac{(\log\log n)^{l-1}}{(l-1)!\log n}, \quad \text{if} \quad l \neq 0, \qquad (n \to \infty) \tag{141}$$

is implied by its truth in the particular case $l = 1$, that is, by the prime number theorem. But (139) can be reduced to (141).

In fact, if $l \neq 0$ is fixed, it is clear from the definitions of $S_l$ and $R_l$ that a positive integer is in $S_l$ if and only if it is a product of the form $ij$, where $i$ is contained in $R_l$, say $i = p_1 \cdots p_l$, and $j$ has *only multiple* prime factors, each of which is distinct from all of the $l$ distinct primes $p_1, \cdots, p_l$. Hence, in order to prove that (141) implies (139), where $l$ is fixed, it is sufficient to show that $\Sigma h^{-1} = A$, where $h$ runs through the sequence of *all* those positive integers

none of which possesses any simple prime factor. In other words, it is sufficient to show that, if $1 = h_1 < h_2 < \cdots$ is the sequence defined before (117), then the series $\Sigma\, h_n^{-1}$ converges to the value (138). But this follows by placing $s = 1$ in (117), where $\sigma > \frac{1}{2}$.

This proves (139). In the limiting case, $l = 0$, the following refinement of (140) is of the same depth as the prime number theorem, that is, as (139) itself.

**64.** Let $s$ be replaced by $s - \frac{1}{2}$ in the identity (117), which is valid for $\sigma > \frac{1}{2}$. Then, if

$$g(n) = n^{\frac{1}{2}}\theta_0(n), \tag{142}$$

(117) appears in the form $\Sigma\, g(n)n^{-s} = \eta(s)\zeta(2s - 1)$, where $\sigma > 1$ and

$$\eta(s) = \zeta(3s - \tfrac{3}{2})/\zeta(6s - 3). \tag{143}$$

But $\Sigma\, g(n)n^{-s} = \eta(s)\zeta(2s - 1)$ means, by (25), that

$$\sum g'(n)n^{-s} = \eta(s)\zeta(2s - 1)/\zeta(s). \tag{144}$$

Since the Dirichlet series of the function (143) is absolutely convergent at $s = 1$, and since the Dirichlet series of the function $\zeta(2s - 1)/\zeta(s)$ was seen to be convergent at $s = 1$ (§14), it follows from the multiplication theorem of Stieltjes, that the Dirichlet series (144) is convergent at $s = 1$. According to Dirichlet's analogue of Abel's continuity theorem, the value represented by the series (144) at $s = 1$ can be obtained by letting $s \to 1 + 0$ in (144). Then $\zeta(2s - 1)/\zeta(s) \to \frac{1}{2}$ and so, from (143) and (144),

$$\sum g'(n)/n = \tfrac{1}{2}\text{ const.}, \qquad \text{where const.} = \zeta(\tfrac{3}{2})/\zeta(3). \tag{145}$$

The relation (140) is an elementary corollary of the variant (145) of the prime number theorem (and can be proved elementarily, cf. Erdös and Szekeres [21], where an $O$-estimate of the remainder of (140), but not of course the corresponding $o$-estimate resulting from (145), is obtained; the numerical value of the constant occurring in (140), a value which also follows from an identity of Cantor [6], is loc. cit. [21] not given in the above form).

First, since the function (142) is real and non-negative, the convergence of $\Sigma\, g'(n)/n$ implies, by ($\mathrm{IX}_2$), the existence of $M(g)$. But (I) shows that $M(g)$ has precisely the value (145). Thus, if (142) is substituted into (6),

$$\sum_{m=1}^{n} m^{\frac{1}{2}}\theta_0(m) = \tfrac{1}{2}\text{ const.}n + o(n); \quad \text{whence} \quad \sum_{m=1}^{n} \theta_0(m) = \text{const.}n^{\frac{1}{2}} + o(n^{\frac{1}{2}}),$$

by partial summation. According to (136), the last relation is identical with (140).

**65.** The characteristic functions $\theta_0(n)$, $\theta_1(n)$, $\cdots$ of the sets $S_0$, $S_1$, $\cdots$ introduced in §62 define a classification of all positive integers into a sequence of mutually disjoint subsets, $S_m$. A dual of the corresponding asymptotic prob-

lem results if the characteristic functions $\theta_0(n)$, $\theta_1(n)$, $\cdots$ are replaced by the functions $\beta_1(n), \beta_2(n), \cdots$ which are defined as follows: $\beta_m(n)$ is the non-negative integer representing the number of those positive integers among the exponents $a, b, \cdots$ of $n = p^a q^b \cdots$ which happen to have the value $m$, where it is understood that $p, q, \cdots$ are distinct primes and that $\beta_m(1) = 0$. In other words, $\beta_m(n)$ denotes the number ($\geqq 0$) of the $m$-fold prime factors of $n$, where $m = 1, 2, \cdots$. Thus (15) shows that, for every fixed $m$,

$$(146) \qquad \beta_m'(p^k) = \delta_{km} m \quad \text{and} \quad \beta_m'(n) = 0 \quad \text{unless} \quad n = p^k, \qquad (k = 1, 2, \cdots),$$

where the notation is the same as in (121).

If $\omega(n)$ denotes, as in §36, the number of all prime divisors of $n$, then obviously

$$(147) \qquad \omega(n) = \sum_{m=1}^{\infty} \beta_m(n) \quad \text{for every } n.$$

It is easily verified from (146) and (147) that everything that was proved in §36 for the function $f(n) = \omega(n) - \nu(n)$ holds for the function $f(n) = \omega(n) - \beta_1(n)$ also. In particular, the latter function is almost periodic $(B)$; so that, since $M(\omega)$ does not exist, $M(\beta_1)$ cannot exist.

The situation is quite different if $m = 1$ is replaced by an arbitrary $m \neq 1$. In fact, although $\beta_m(n) \neq O(1)$ as $n \to \infty$ holds for every $m$, the function $\beta_m(n)$ of $n$ is almost periodic $(B)$, with a Fourier series $(B)$ representing $\beta_m(n)$ for every $n$, if $m \neq 1$. This follows by observing that, according to (146),

$$\sum_{n=1}^{\infty} d(n)\beta_m'(n)/n = \sum_p d(p^m)m/p^m = (m + 1)m \sum_p 1/p^m,$$

and so the assumption of (XVIII) is satisfied unless $m = 1$.

Almost periodicity of any kind is more than sufficient for the existence of an asymptotic distribution function (as to the latter notion, cf. Jessen and Wintner [48], Hartman, van Kampen and Wintner [44], Erdös and Wintner [22]). Since $\beta_m(n)$ is capable of the values $0, 1, 2, \cdots$ only, it follows that there exist limits $\alpha_0^m, \alpha_1^m, \cdots$ satisfying

$$(148) \qquad \lim_{n\to\infty} S_l^m(n)/n = \alpha_l^m \quad \text{and} \quad \sum_{l=0}^{\infty} \alpha_l^m = 1, \quad \text{if} \quad m \neq 1,$$

where $S_l^m$ denotes the set of those positive integers $n$ for which $\beta_m(n) = l$, and $S_l^m(n)$ is defined by (136).

According to the definition of $\beta_m(n)$, the function $\beta_1(n)$ represents the number of the simple prime factors of $n$. Hence the preceding definition shows that

$$(149) \qquad S_l^1 = S_l \quad \text{for} \quad l = 0, 1, 2, \cdots,$$

where $S_l$ is the set defined after (136). Thus (139) and (140) represent what corresponds to (148) in the case $m = 1$.

**66.** The interpretation (149) of $S_l$ leads to a general lemma, which may be formulated as follows (cf. Wintner [94]):

Let $T$ be any set of positive integers such that there exists a positive integer $l = l_T$ for which (i) no integer contained in any of the sets $S_{l+1}, S_{l+2}, \cdots$ is in $T$ and (ii) an integer contained in $S_l$ is or is not in $T$ according as it is or is not square-free. Then, in terms of the notation (136),

$$\frac{T(n)}{n} \sim \frac{(\lambda(n))^{l-1}}{(l-1)!} e^{-\lambda(n)} \quad \text{as} \quad n \to \infty, \tag{150}$$

where

$$\lambda(n) = \log \log n. \tag{151}$$

In order to prove this, let $L(n)$ be an abbreviation for the expression on the right of (150), where $l$ is fixed. Then, from ($139_0$) and (139),

$$S_0(n)/n + S_1(n)/n + \cdots + S_{l-1}(n)/n = o(L(n)). \tag{152}$$

Those integers contained in $T$ for which neither (i) nor (ii) assumes anything are contained in the set $S_0 + S_1 + \cdots + S_{l-1}$, and so their relative frequency is majorized by (152). On the other hand, the relative frequency of those integers contained in $T$ which are contained in the set $S_{l+1} + S_{l+2} + \cdots$ is 0, by assumption (i). Hence, in order to prove (150), it is sufficient to show that the relative frequency of those integers contained in $T$ which are contained in $S_l$ is $L(n) + o(L(n))$. But since $S_l$ consists of the positive integers having exactly $l$ simple prime factors, assumption (ii) can be expressed by saying that an integer contained in $S_l$ is contained in $T$ if and only if it is square-free. Consequently, the assertion is that the relative frequency of the integers composed of exactly $l$ distinct primes is $L(n) + o(L(n))$, where $L(n)$ denotes the function on the right of (150). Since (151) shows that the expression on the right of (141) is identical with $L(n)$, and since $R_l(n)/n$ denotes precisely the relative frequency just mentioned, it follows from (141) that the assertion is true. This proves (150).

**67.** Corresponding to the definition, (37), of the class of multiplicative functions, a function $f(n)$ is called additive if

$$f(n_1 n_2) = f(n_1) + f(n_2) \quad \text{whenever} \quad (n_1, n_2) = 1. \tag{153}$$

Since this means that

$$f(1) = 0 \quad \text{and} \quad f(n) = f(p^k) + f(q^j) + \cdots \quad \text{if} \quad n = p^k q^j \cdots, \tag{154}$$

where $p, q, \cdots$ are distinct primes, it is clear from (15) that $f(n)$ is additive if and only if

$$f'(n) = 0 \quad \text{unless} \quad n = p^k, \quad (k = 1, 2, \cdots; p = 2, 3, \cdots), \tag{155}$$

as illustrated by (121) and (146). Accordingly, an additive function is uniquely determined by an arbitrary assignment of a double sequence of values $f(p^k)$. Since

$$f(1) = 0, \quad \text{hence} \quad f'(1) = 0 \tag{156}$$

in virtue of (18), it is clear from (17) that such an arbitrary assignment is equivalent to an arbitrary assignment of the values

$$(157)\quad f'(p) = f(p), \qquad f'(p^2) = f(p^2) - f(p), \qquad f'(p^3) = f(p^3) - f(p^2), \cdots$$

for every $p$.

For every fixed prime $q$ and for every additive function $f = f(n)$, let $f_q = f_q(n)$ denote the additive function assigned by the values

$$(158)\qquad f_q(p^k) = \delta_{pq} f(p^k), \qquad (k = 1, 2, \cdots ; p = 2, 3, \cdots),$$

where $\delta_{pq} = 0$ if $p \neq q$ and $\delta_{qq} = 1$. Thus it is clear that, corresponding to the representation (43), (45) of a multiplicative function,

$$(159)\qquad f(n) = \sum_p f_p(n) \text{ for every } n$$

(it being understood that the infinite sum (159) has only a finite number of non-vanishing terms for every fixed $n$).

It is known that, if $f(n)$ is additive, then

($I$) $f(n)$ has an asymptotic distribution function if and only if both series $\Sigma f^*(p)/p$, $\Sigma \mid f^*(p) \mid ^2/p$ are convergent, where $f^*(p)$ denotes $f(p)$ or 1 according as $\mid f(p) \mid < 1$ or $\mid f(p) \mid \geqq 1$ (Erdös and Wintner [22]);

($II$) $f(n)$ is almost periodic ($B$) if and only if all four series

$$\sum_p \frac{f(p)}{p}, \qquad \sum_{k=2}^{\infty} \sum_p \frac{\mid f(p^k) \mid}{p^k}, \qquad \sum_{\mid f(p) \mid \geqq 1} \frac{\mid f(p) \mid}{p}, \qquad \sum_p \frac{\mid f^*(p) \mid^2}{p}$$

are convergent (Hartman and Wintner [42]); a criterion obtained by an adaptation of the proof of the more elegant result according to which

($III$) $f(n)$ is almost periodic ($B^2$) if and only if both series $\Sigma f(p)/p$, $\Sigma\Sigma \mid f(p^k) \mid ^2/p^k$, where $k = 1, 2, \cdots$, are convergent (Erdös and Wintner [23]);

($IV$) $f(n)$ is almost periodic ($W$) if and only if $f(n) = O(1)$ as $n \to \infty$ (Hartman and Wintner [43]);

($V$) $f(n)$ is uniformly almost periodic if and only if the series (159) is uniformly convergent for $n = 1, 2, \cdots$ (this follows from the criterion of Toeplitz [79], quoted at the end of §51; cf. van Kampen [50]).

These criteria, the first three of which lie quite deep (even though they do not involve the prime number theorem), and none of which has an analogue in case of an arbitrary function $f(n)$, will not be proved here; they are not needed in the sequel.

**68.** For an arbitrary function $f(n)$ and for a given value $a$, let $F_a$ denote the set of those positive integers $n = n_a$ which is defined by

$$(160)\qquad F_a: \qquad f(n_a) = a.$$

Thus, if $F_a(n)$ is the function (136) belonging to $S = F_a$, then $F_a(n)/n$ is the relative frequency of the value $a$ among the $n$ values $f(1), \cdots, f(n)$. In particular, the set $F_a$ is vacuous unless $f(n)$ attains the value $a$ for some $n$.

In view of (156), an additive function will be called positive if $f(n) > 0$ for every $n \neq 1$. This means that $F_a$ is vacuous unless $a \geqq 0$ and that

$$F_0(n) = 1 \quad \text{for every } n; \tag{161}$$

or, what by (154) is the same thing, that each of the values $f(p^k)$ is positive.

It will now be shown that the lemma of §66 implies the following theorem (Wintner [90], [94]):

(†) *If $f(n)$ is a positive additive function satisfying $f(p) = 1$ for every prime $p$, then, no matter what the* (*positive*) *values $f(p^2), f(p^3), \cdots$ assigned for $p = 2, 3, \cdots$ may be, the asymptotic distribution of $f(n)$ over the sets $F_1, F_2, \cdots$ obeys Poisson's statistical law of independent rare events*, that is,*

$$\frac{F_l(n)}{n} \sim \frac{(\lambda(n))^{l-1}}{(l-1)!} e^{-\lambda(n)} \quad \text{as} \quad n \to \infty, \qquad (l = 1, 2, \cdots), \tag{162}$$

*where $\lambda(n)$, the squared standard deviation* (*variance*), *is* $\log \log n$.

The occurrence of the function (151) in (162) is a manifestation of Mertens' elementary result, used in §57. On the other hand, (162) depends on the prime number theorem (which it contains for $l = 1$).

Since the values $f(p^k)$ assigned for $k > 1$ need not be integers, it is clear that what corresponds to ($137_1$) is

$$1 + \sum_{l=1}^{\infty} F_l(n) + F^*(n) = n \quad \text{for every } n, \tag{163}$$

where the first term represents the contribution, (161), of the set $F_0$, and $F^*$ denotes the set of those positive integers at which the value attained by $f$ is not an integer. However, it is easy to adapt the proof given in §66 so as to yield the estimate

$$F^*(n) = o(n); \qquad \text{hence} \qquad \sum_{l=1}^{\infty} F_l(n)/n \to 1 \tag{164}$$

holds in virtue of (163). Since, from (151),

$$\sum_{l=1}^{\infty} \frac{(\lambda(n))^{l-1}}{(l-1)!} e^{-\lambda(n)} = \sum_{l=0}^{\infty} \frac{(\log \log n)^l}{l!} \frac{1}{\log n} = \frac{e^{\log \log n}}{\log n} = 1, \tag{165}$$

it follows that the sequence of the asymptotic laws (162) represents a case of *complete additivity.*

It is by no means clear that this *must* be the case. In fact, the situation is quite different for either of the classifications represented by (141) and (139)-(140); classifications in which the sum of the elementary approximations is 1 and $A$ respectively. In fact, the classification (141) refers to all square-free integers, and the latter are known to have the asymptotic relative frequency $1/\zeta(2)$, which is distinct from 1. On the other hand, the classification considered by

* Cf., e.g., G. Darmois, Statistique Mathématique, Paris, 1928, p. 84.

(139)-(140) enumerates the set of all positive integers, and the latter have the frequency 1, which, by (138), is distinct from $A$.

Incidentally, (138)-(139) also show that the assumptions, $f(p) = 1$ and $f(p^k) > 0$, of (162) are so essential that (162) becomes false if only $f(p) = 1$ and $f(p^k) \geqq 0$ are assumed. In fact, if $f(n) = \beta_1(n)$, where $\beta_1(n)$ is the number of the simple prime factors of $n$, then the set (160) belonging to $a = l$ is precisely the set (149) to which (138)-(139) refer. Nevertheless, $f(n) = \beta_1(n)$ is a non-negative additive function satisfying $f(p) = 1$ (however, $f(p^2) = f(p^3) = \cdots = 0$). Needless to say, the assumption $f(p) = 1$ of (162) is only a normalization of the assumption $f(p) =$ const. $> 0$. If const. $= 0$, it is illustrated by (148) how radically is the situation changed.

**69.** The proof of the Poissonian theorem (†), §68 proceeds as follows:

Let $l$ be a fixed positive integer. Since $f(p) = 1$ and $f(p^k) > 0$ by assumption, it is clear from (154) that $f(n)$ is equal to or greater than the number of the simple prime factors of $n$, according as $n$ is or is not square-free. Hence, if $S_m$ denotes the set of those positive integers having exactly $m$ simple prime factors (a set containing not *only* square-free integers), then (i) no $n$ contained in $S_{l+1}, S_{l+2}, \cdots$ can satisfy $f(n) = l$ and (ii) an $n$ contained in $S_l$ satisfies $f(n) = l$ if and only if it is square-free. Hence both assumptions, (i)-(ii), of §66 are fulfilled by the set $T$ consisting of those positive integers $n$ for which $f(n) = l$, where $l$ is a fixed positive integer. Since $T = F_l$ in virtue of (160), it follows from (150) that the proof of (162) is complete.

**70.** It is worth pointing out the common formal background of (†), §68 and of $(I)$-$(V)$, §67.

Let $g_1(n), \cdots, g_m(n)$ be $m$ periodic functions of $n$, and suppose that their periods, say $k_1, \cdots, k_m$, are relatively prime (in the narrower sense of the term, that is, in the sense that $(k_i, k_j) = 1$ unless $i = j$). Then, since the product $k = k_1 \cdots k_m$ is a common period of $g_1, \cdots, g_m$, a straightforward counting (or, equivalently, a verification of "ergodicity" with reference to an "incompressible flow") shows that

$$(166) \qquad \sum_{n=1}^{k} \prod_{j=1}^{m} (g_j(n))^{r_j} = \prod_{j=1}^{m} \sum_{n=1}^{k_j} (g_j(n))^{r_j}, \quad \text{where} \quad k = \prod_{j=1}^{m} k_j,$$

is an identity in the $m$ non-negative integers $r_1, \cdots, r_m$. If the functions are real-valued, then (166) means that $g_1(n), \cdots, g_m(n)$ are statistically independent (in the sense in which this term has always been used*). The complications arising in case of complex-valued functions are only typographical in nature.

Now let $f(n)$ be an additive function and suppose, for a moment, that there exists for every $p$ a sufficiently large $m = m_p$ so that $f'(p^{m+1}) = f'(p^{m+2}) = \cdots 0$. Then it is clear from (15) and (158) that each of the additive functions $f_p(n)$

* Cf., e.g., p. 194 of Darmois's book, referred to in the preceding footnote, or A. Kolmogoroff's *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933, p. 50.

occurring in (159) is periodic; the primitive period of $f_p(n)$ being $p^m$, if $m = m_p$ has its least value. It follows that the functions $g_1(n) = f_p(n)$, $g_2(n) = f_q(n)$, $\cdots$ representing the various terms of the series (159) have periods $k_1$, $k_2$, $\cdots$ satisfying the assumption of (166).

If $m = m_p$ does not exist, $f_p(n)$ is not strictly periodic. However, the statistical independence still holds in terms of asymptotic distributions (cf. Wintner [88]; Hartman, van Kampen and Wintner [44]); as a matter of fact, it holds (cf. Hartman and Wintner [43], §3) even in the so-called unrestricted sense (cf. Jessen and Wintner [48], §11).

Incidentally, all of this can be extended to the more general case indicated at the end of the footnote to (21).

## The Statistics of the Sum of Two Squares

**71.** Let $g(n)$ denote the function for which the function $g'(n)$, defined by (15), attains the value $(-1)^{\frac{1}{2}(n-1)}$ or 0 according as $n$ is odd or even. Then obviously

$$(167) \qquad g(2^k) = 1, \qquad g(q^k) = k + 1, \qquad g(r^k) = \tfrac{1}{2} + \tfrac{1}{2}(-1)^k$$

for $k = 1, 2, \cdots$, if $q$ and $r$ denote those primes for which

$$(168) \qquad q \equiv 1 \pmod 4 \quad \text{and} \quad r \equiv 3 \pmod 4)$$

respectively (it is understood that $\frac{1}{2} + \frac{1}{2}(-1)^k$ in (167) denotes 1 or 0 according as $k$ is even or odd).

Since the function $g'(n)$ is multiplicative, the same is true, by (39), of the function $g(n)$. If $g'(n)$ is replaced by $f'(n) = 4g'(n)$, then the sum (15) becomes the classical formula for the number, $r_2(n)$, of the representations of $n$ as a sum of two squares. In other words, $g(n)$ denotes *one quarter* of the number of those points in a planar square lattice of unit width which have the distance $n^{\frac{1}{2}}$ from the origin (0, 0) of the lattice.

In accordance with the notation (160), let all positive integers $n$ be classified into a sequence of mutually disjoint sets $G_0$, $G_1$, $\cdots$, by placing $n$ into $G_l$ if and only if $g(n) = l$. Since $g(n)$ is a multiplicative function, it is clear from (167) that every "urn" $G_l$ contains an infinity of integers.

The majority of the integers $n$ are in $G_0$, in the sense that, in terms of the notation (136),

$$(169) \qquad G_0(n)/n \to 1 \quad \text{as} \quad n \to \infty$$

(for a sharper result, cf. (170)-(171) below). If $\gamma(x)$ denotes the asymptotic relative frequency ("probability") of the inequality $g(n) < x$, where $x$ is a fixed real number, then (169) is equivalent to the statement that $\gamma(x) = 0$ for $-\infty < x < 0$ and $\gamma(x) = 1$ for $0 < x < \infty$. This implies that $g(n)$ cannot be almost periodic $(B)$. For if it were, $M(g)$ ought to be represented by the first moment of the asymptotic distribution function, $\gamma$, of $g$, that is, by the Stieltjes integral $\int x d\gamma(x)$ over the range $-\infty < x < \infty$. Since this integral is 0, there results a contradiction with the evaluation $M(g) = \pi/4$, obtained in §7 as a consequence

of $(29_2)$. [However, if $a_m$ denotes $(-1)^{\frac{1}{2}m-\frac{1}{2}}\pi/m$ or 0 according as $m$ is odd or even, then the trigonometric series (86) represents $f(n) = 4g(n)$ for every $n$ (Ramanujan [69], p. 190; the convergence of the series is substantially equivalent to the prime number theorem).]

**72.** There arises the question as to the asymptotic law (if any) of the relative frequencies, $G_l(n)/n$, determined by the distribution of all positive integers over the various "urns" $G_l$, where $l = 0, 1, \cdots$. The answer,* which will be given completely, turns out to be quite involved. In view of everything that is known on questions of this type, it is not surprising that the problem happens to depend on the machinery of the prime number theorem. What makes the situation involved is not this methodical circumstance but rather the fact that the order of magnitude of the relative frequency $G_l(n)/n$ for large $n$ and fixed $l$ depends on the *arithmetical structure of* $l$ in an intricate manner, instead of being dependent, as in the Poissonian law (†), §68, on the mere size of $l$. Correspondingly, the result to be obtained disposes not only of the crude mechanism considered by Tricomi [81] but also of the possibility of the statistical model subsequently suggested by Lévy [61].

Clearly, the analogue of (163) is

$$(170) \qquad G_0(n) + G^\infty(n) = n, \quad \text{if} \quad G^\infty(n) = \sum_{l=1}^{\infty} G_l(n),$$

that is, if $G^\infty$ denotes the set, $G_1 + G_2 + \cdots$, of those positive integers which can be represented in at least one way as a sum of two squares. It is known that (169) can be refined to

$$(171) \qquad G^\infty(n)/n \sim C^\infty(\log n)^{-\frac{1}{2}}, \qquad \left(G^\infty = \sum_{l=1}^{\infty} G_l\right),$$

where $C^\infty$ (as every $C$ in the sequel) denotes a positive numerical constant (Landau [57]; cf. also Watson [84], Wintner [93]). The question formulated before refers to a corresponding asymptotic relation (if any) for each of the terms of the infinite sum (170).

It will be shown that

$$(172) \qquad G_1(n)/n \sim C_1(n \log n)^{-\frac{1}{2}}$$

(which, in view of (171), implies that the first term of the infinite sum (170) is of a much lower order than the sum itself); that

$$(173) \qquad G_2(n)/n \sim C_2(\log n)^{-1}$$

* The question can be interpreted in terms of the distribution of the characteristic values of a vibrating quadratic membrane (cf., e.g., Riemann-Hattendorff, *Partielle Differentialgleichungen*, 2nd ed. (1876), pp. 250–258), but the answer can hardly have an acoustical interest.

(so that the second term is almost of the same order as the sum itself); that

(174) $$G_3(n)/n \sim C_3(n \log n)^{-\frac{1}{2}} \log \log n$$

(so that the third term is of almost the same order as the first); that

(175) $$G_l(n)/n \sim C_l(n \log n)^{-\frac{1}{2}} \quad \text{if} \quad l = p^k > 3, \quad (k = 1, 2, \cdots),$$

(so that $l = 2$ and $l = 3$ are the only prime powers for which the order of $G_l(n)$ is not the same as that of $G_1(n)/n$); finally, that one of the three possibilities

(176) $$G_l(n) \sim \text{const. } G_\lambda(n); \qquad \lambda = 1, 2, 3, \qquad (\text{const.} > 0)$$

must take place for every positive integer $l$ (the actual value of $\lambda$ in (176) is an arithmetical function of $l$).

**73.** Let $g_l(n)$ denote the characteristic function of the set $G_l$. Then, corresponding to (136),

(177) $$G_l(n) = \sum_{m=1}^{n} g_l(m).$$

The generating Dirichlet series is

(178) $$\phi_l(s) = \sum g_l(n) n^{-s}.$$

It will be necessary to consider the function (178) in its dependence on the arithmetical structure of $l$.

In view of (170) and (171), it can be assumed that $l \neq 0$, that is, that $n$ is in one of the sets $G_1, G_2, \cdots$. This disposes of all but those integers $n$ which are of the form

(179) $$n = 2^h \prod q^j \prod r^{2i},$$

where $h$, $j$, $i$ denote non-negative integers and $q$, $r$ represent the two classes (mod 4) of odd primes, as defined by (168). In fact, since $g(n)$ is a multiplicative function, it is clear from (167) that (179) is necessary and sufficient for $g(n) \neq 0$, that is, for an $n$ not contained in $G_0$.

The coefficient $g_l(n)$ of (178) is 1 or 0 according as $g(n) = l$ does or does not hold for the multiplicative function $g(n)$ defined by (167). Hence it is clear from (167) and (179) that (178) is reduced for $l = 1$ to

(180) $$\phi_1(s) = \sum (2^h \prod r^{2i})^{-s}, \qquad (\sigma > 1).$$

Next, if $l$ has any fixed value which is a prime power, say $l = p^k$, where $k = 1, 2, \cdots$, then it is clear that the multiplicative function $g(n)$ defined by (167) satisfies $g(n) = l$ only if the first product in the factorization (179) of $n$ is reduced to a single factor, $q^j$; and that this necessary condition becomes sufficient if and only if the exponent of $q^j$ is restricted to have the fixed value assigned by $j = l - 1$. This means that (178) is reduced to

$$\phi_l(s) = \sum (2^h q^{l-1} \prod r^{2i})^{-s}, \quad \text{if } l = p^k.$$

Hence, if $\omega(s)$ is an abbreviation for

(181) $$\omega(s) = \sum q^{-s}, \qquad (\sigma > 1),$$

it is seen from (180) that

(182) $$\phi_l(s) = \phi_1(s)\omega([l - 1]s), \quad \text{if} \quad l = p^k.$$

It is clear from this deduction that, if $l$ is the product of two prime powers, say of $a$ and of $b$, where $(a, b) = 1$, then (182) must be replaced by

(183) $$\phi_{ab}(s) = \phi_1(s) \sum q_1^{-(a-1)s} q_2^{-(b-1)s},$$

where the summation ranges over all distinct pairs of distinct indices $q_1$, $q_2$. It is understood that a $q$ always denotes a prime satisfying the first of the conditions (168).

Finally, if $l = abc \cdots$, where $a, b, c, \cdots$ are powers of distinct primes, then $\phi_l(s)$ results if the double sum on the right of (183) is replaced by the corresponding multiple sum.

**74.** Clearly, the Dirichlet series (180) has the Euler factorization

$$\phi_1(s) = \sum_{h=0}^{\infty} 2^{-hs} \prod_r \sum_{i=0}^{\infty} r^{-2is} = (1 - 2^{-s})^{-1} \prod_r (1 - r^{-2s})^{-1}.$$

Hence, from $\zeta(s) = \Pi(1 - p^{-s})^{-1}$, where $p = 2, q, r$ in the notations of (168),

(184) $$\phi_1(s) = (1 + 2^{-s})\zeta(2s)\eta(2s),$$

where, as an abbreviation,

(185) $$\eta(s) = \prod (1 - q^{-s}), \qquad (\sigma > 1).$$

Thus the behavior of all the Dirichlet series (178) on the boundaries of their respective half-planes of convergence is substantially reduced to the behavior of $\zeta(s)$ and of $\eta(s)$ on the line $\sigma = 1$.

It is known from the theory of the $L$-series of Dirichlet that, if $q$ runs through all primes satisfying $q \equiv k \pmod m$, where $(k, m) = 1$, then the contribution of (185) to the behavior of $1/\zeta(s) = \Pi(1 - p^{-s})$ on the line $\sigma = 1$ is proportional to the ratio $\alpha = 1 : \phi(m)$; in the sense that the function $(s - 1)^{-\alpha}\eta(s)$ is regular and distinct from 0 on the line $\sigma = 1$ (for sharper results, cf. Watson [84]). Since $q \equiv 1 \pmod 4$ in the present case, it follows from $\phi(4) = 2$ that

(186) $$\eta(s) = c_0(s - 1)^{\frac{1}{2}} + \rho_0(s) \quad \text{for} \quad \sigma \geqq 1,$$

and so, from (184), that

(187) $$\phi_1(s) = c_1(s - \tfrac{1}{2})^{-\frac{1}{2}} + \rho_1(s) \quad \text{for} \quad \sigma \geqq \tfrac{1}{2},$$

where, as always in the sequel, every $c$ is a positive constant and every $\rho(s)$ denotes a function satisfying the following conditions: It is regular in the interior

of the half-plane specified and possesses a continuous boundary function. For instance, from (181) and (185),

$$\omega(s) = -\log \eta(s) + \rho(s) \quad \text{for} \quad \sigma \geqq 1. \tag{188}$$

**75.** It is known that Ikehara's theorem can be extended as follows (cf. Wintner [93]):

If a Dirichlet series $\Sigma f(n)n^{-s}$, where $f(n) \geqq 0$, converges for $\sigma > 1$ and is such that there exist two positive constants, $B$ and $\beta$, for which the function $\Sigma f(n)n^{-s} - B/(s-1)^{\beta}$ acquires continuous boundary values on the line $\sigma = 1$, then

$$\sum_{m=1}^{n} f(m) \sim \text{const.}\, n/(\log n)^{1-\beta} \quad \text{as} \quad n \to \infty, \quad \text{where const.} = B/\Gamma(\beta). \tag{189}$$

The logarithm disappears in Ikehara's case, where $\beta = 1$.

Since every $g_l(n)$ in (178) is either 1 or 0, it follows from (187) that the case $\beta = \frac{1}{2}$ of the theorem is applicable to the Dirichlet series of the function $\phi_1(s + \frac{1}{2})$. In fact, the replacement of $s$ by $s - \frac{1}{2}$ translates the line $\sigma = \frac{1}{2}$ of (187) into the line $\sigma = 1$ of the theorem, and transforms the coefficient $g_1(n)$ into $g_1(n)n^{\frac{1}{2}}$, a coefficient which is again real and non-negative. Accordingly,

$$\sum_{m=1}^{n} g_1(m)m^{\frac{1}{2}} \sim \text{const.}\ (n/\log n)^{\frac{1}{2}}, \qquad (\text{const.} > 0). \tag{190}$$

It is seen from (177) by partial summation that (190) completes the proof of (172).

In the case $l = p^k$ of (182), it is clear from (188) that, unless $l - 1 \leqq 2$, the "nearest" singularity of the second factor of (182) lies beyond the critical line, $\sigma = \frac{1}{2}$, of the first factor, (187). Hence, (175) follows in exactly the same way as (172). Thus, if $l = p^k$, only the cases $l - 1 \leqq 2$, that is $l = 2$ and $l = 3$, remain to be treated.

If $l = 2$, the second factor on the right of (182) is the function (188). Hence it is clear from (186) and (187) that (182) is now of the form

$$\phi_2(s) = -c_2 \log(s-1) + \rho_2(s) \quad \text{for} \quad \sigma \geqq 1. \tag{191}$$

Consequently, (173) follows by comparing the case $l = 2$ of (177) and (178) with

$$\sum_{p<n} 1 \sim n/\log n \quad \text{and} \quad \sum_{p} 1 \cdot p^{-s} = \log \zeta(s) + \rho(s), \quad \text{where} \quad \sigma \geqq 1,$$

and observing that the analytic behavior of $\log \zeta(s)$ on the line $\sigma = 1$ is the same as that of the function (191).

If $l = 3$, then (188), (187), (186) and (182) show that (191) must be replaced by

$$\phi_3(s + \tfrac{1}{2}) = -c_3(s - \tfrac{1}{2})^{-\frac{1}{2}} \log(s - \tfrac{1}{2}) + \rho_3(s) \quad \text{for} \quad \sigma \geqq \tfrac{1}{2}. \tag{192}$$

It follows that

$$\sum_{m=1}^{n} g_3(m)m^{\frac{1}{2}} \sim \text{const.}(n \log\log n)/(\log n)^{\frac{1}{2}}, \qquad (\text{const.} > 0). \tag{193}$$

In fact, the order assigned by (192) for the case $l = 3$ of (178) is an Abelian consequence of (193). On the other hand, the Tauberian transition from (192) to (193) follows from the general theorem which results by rewriting, for the case of the singularity of (192), the general theorem used in the Tauberian transition from (187) to (190). Correspondingly, the partial summation leading from (190) to (172) leads from (193) to (174).

This completes the treatment of the cases $l = 1$ and $l = p^k$.

**76.** Suppose now that $l = ab$, where $a$ and $b$ are of the form $p^k$, and $(a, b) = 1$. Then $\phi_l(s)$ is given by (183). But it is seen from (181) that (183) is substantially the same function as

$$\phi_1(s)\omega([a - 1]s)\omega([b - 1]s); \tag{194}$$

the necessary additive correction being just the function

$$\sum q^{-(a-1)s} q^{-(b-1)s} = \omega([a + b - 2]s), \tag{195}$$

the effect of which can be appraised as follows:

First, $a + b \geqq 2 + 3$, since $a$ and $b$ are either distinct primes or powers of such primes. Hence it is clear from (181) that the additive correction, (195), is regular in the half-plane $3\sigma > 1$. Since this half-plane contains not only the half-plane $\sigma \geqq 1$ but the half-plane $\sigma \geqq \frac{1}{2}$ as well, it follows from (187) and (188) that the relevant "nearest singularities" of (183) can always be replaced by those of (194).

As in the case, $l = p^k$, of §75, three subcases of the present case, $l = ab$, will have to be distinguished.

The least possible value of $a$ is 2. Suppose first that $a = 2$ in $l = ab$. Then $b \geqq 3$. Hence it is clear from (187) and (188) that the analytic behavior of (194), and therefore that of (183), on the critical line is the same as that of (191). Consequently, (176) holds for $\lambda = 2$.

Next, suppose that $a = 3$ in $l = ab$. Thus, since the case $a = 3$, $b = 2$ is identical with the case $a = 2$, $b = 3$, considered before, it can be assumed that $b = p^k$ is at least 4. Hence, for the same reasons as in the preceding case, it now follows that (176) holds for $\lambda = 3$.

Finally, the same reasoning shows that (176) holds for $\lambda = 1$ in the remaining case of the assumption $l = ab$, that is, in case $a$ and $b$ are powers of two distinct primes both of which exceed 3.

Since all of this can be transcribed from $l = ab$ to the general case $l = abc \cdots$, described at the end of §73, the proof of the assertions of §72 is complete.

## The Ergodic Law of Multiplicative Sets

**77.** Let a set, $S$, of positive integers be called completely multiplicative if it not only is multiplicative in the sense of §52 but is such as to contain the product of any two of its elements, even if the latter are not relatively prime (or distinct). Thus $S$ is uniquely determined by its multiplicative basis, that is, by the se-

quence, say $S^\dagger$, of the primes contained in $S$. Conversely, any given sequence of primes generates a completely multiplicative set.

The prime number theorem and the results connected with it succeed, to some extent, in analyzing the unknown set $S^\dagger$ in terms of the given set $S$ when the latter is the sequence of all positive integers. The success of this analysis depends on that esoteric state of saturation which compels the set $S$ to be equidistant when $S^\dagger$ is given as the sequence of "all" primes. In fact, the elementary approximations of Chebyshev and Mertens to the rough order of magnitude of the $m$-th prime involve, via Stirling's theorem, the explicit assumption that $S$ be equidistant, and supply explicit statements as to the asymptotic distribution of $S$.

However, there exist less explicit forms of the prime number theorem; formulations which do not refer at all to the order of magnitude of the $m$-th prime. For instance, it is known that the prime number theorem is equivalent to the assertion that the mean, $M(\mu)$, of the Möbius function, $\mu(n)$, exists; (Landau [58]; cf. Hardy and Littlewood [37]). Since the value of $M(\mu)$ is 0, and since $\mu(n)$ is capable only of the values 0 and $\pm 1$, the prime number theorem thus appears in an "ergodic" form, stating that there exists an asymptotic probability, $\frac{1}{2}$, that $\mu(n)$ be positive, and therefore the same probability that $\mu(n)$ be negative, if the trivial frequency of the $n$-values for which $\mu(n)$ vanishes is subtracted.

The aim of this chapter is to prove that a corresponding formulation of the prime number theorem holds in case of *any* sequence of primes. In view of the exceptional set of measure 0, excluded by the law of large numbers or, more generally, by Birkhoff's ergodic theorem (an exceptional set which is of the second category in the usual realization of the underlying product measure), it is of course a purely arithmetical fact that the proviso of the "almost all" can be replaced by the assertion that what corresponds to the exceptional set is now *vacuous*.

The results will be such as to hold for arbitrary multiplicative sets, and not only for the completely multiplicative sets defined above.

**78.** With reference to a given multiplicative set, $S$, let the sequence, $P$, of all primes be subdivided into two subsequences, say $Q$ and $R$, by placing a prime into $Q$ or into $R$ according as it is or is not contained in $S$. In order words, if $f(n)$ denotes the multiplicative function representing the characteristic function of $S$, let

$$(196) \qquad p = q \quad \text{when} \quad f(p) = 1 \quad \text{and} \quad p = r \quad \text{when} \quad f(p) = 0,$$

where $q$, $r$ and $p$ denote arbitrary elements of $Q$, $R$ and $P$ respectively. If $k > 1$, then both $f(q^k)$ and $f(r^k)$ can be either 1 or 0 (depending on $k$ and $p$). Since $f(1) = 1$, it follows from (17) that

$$(197) \qquad f'(p^k) = -1, 0, 1, \qquad (k = 1, 2, \cdots),$$

holds in both cases, $p = q$ and $p = r$, and that, if $k = 1$,

$$(198) \qquad f'(q) = 0, \qquad f'(r) = -1.$$

Since $f(n)$ is a multiplicative function, (39) and (197) imply that

$$(199)\qquad f'(n) = -1, 0, 1, \qquad (n = 1, 2, \cdots).$$

It also follows that

$$(200)\qquad \sum_{k=2}^{\infty} \sum_{p} |f'(p^k)|/p^k < \infty, \quad \text{where} \quad f'(p^k) = f(p^k) - f(p^{k-1}).$$

**79.** Correspondingly, the following elementary alternative is a mere restatement of the standard limiting form of the sieve of Eratosthenes:

(i) *If $f(n)$ is the characteristic function of a multiplicative set, $S$, then $M(f)$ exists. Furthermore, $M(f) = 0$ or $0 < M(f) \leqq 1$ according as the sum of the reciprocal values of all primes not contained in $S$ is divergent or convergent; in fact,* (42) *holds in both cases.*

Suppose first that the sum mentioned in (i) is convergent. This means, by (196), that $\Sigma\, r^{-1} < \infty$. It follows therefore from (198) that (200) remains valid if the lower limit, $k = 2$, of the exterior summation is replaced by $k = 1$. Hence the case $g(n) = |f'(n)/n|$ of (51) shows that $\Sigma\, |f'(n)|/n < \infty$. This implies that (50) is true for $g(n) = f'(n)/n$, and that $M(f)$ exists in virtue of the remark following (VIII), §20. Thus (42) is true. Finally, (42) and (196), where $\Sigma\, r^{-1} < \infty$ by assumption, imply that $M(f) > 0$, since every $f(p^k)$ is either 0 or 1.

It is also seen that, if $\Sigma\, r^{-1} = \infty$, the product on the right of (42) is 0. Hence, in order to complete the proof of (i), it is sufficient to show that $M(f)$ exists and vanishes when $\Sigma\, r^{-1} = \infty$.

To this end, let $S_m$ denote, for every positive integer $m$, the multiplicative set which, in terms of its characteristic function, $f_m(n)$, is defined as follows: If $p$ is a $q$, then $f_m(p^k) = f(p^k)$, and if $p$ is an $r$, then $f_m(p^k) = f(p^k)$ or $f_m(p^k) = 1$ according as $r < m$ or $r \geqq m$, where it is understood that $k = 1, 2, \cdots$ in all three cases, and that $q$ and $r$ refer to the disjunction (196). It is clear from this definition that the relations

$$(201)\qquad f_m(n) \leqq f_{m-1}(n) \quad \text{and} \quad \lim_{m\to\infty} f_m(n) = f(n)$$

hold for every $n = p^k$, and therefore for every positive integer $n$. It is also clear that $f_m(q) = 1$ for every $q$ and that $f_m(r) = 0$ or $f_m(r) = 1$ according as $r < m$ or $r \geqq m$.

In particular, $S_m$ contains all but a finite number of primes. Since the reciprocal values of the excluded primes form a finite sum, it follows from what has already been proved that $M(f_m)$ exists and is represented by the product (42) belonging to $f_m(n)$. Hence it is seen from the values $f_m(p^k)$ defining the function $f_m$ (and, in particular, from the values of $f_m(q)$ and $f_m(r)$ pointed out before), that, since $f'(p^k) = f(p^k) - f(p^{k-1})$,

$$M(f_m) = C \prod_{r<m} \left\{1 + \frac{0-1}{r} + \sum_{k=2}^{\infty} \frac{f'(r^k)}{r^k}\right\},$$

where $C > 0$ denotes the value of the product which results if $p$ is replaced by $q$ on the right of (42). Since $C$ is independent of $m$, it follows from (199) that

$$M(f_m) = C \prod_{r<m} \left\{1 - \frac{1}{r} + O \sum_{k=2}^{\infty} \frac{1}{r^k}\right\} \quad \text{as } m \to \infty$$

where the $O$ refers to $r \to \infty$. Hence the assumption $\Sigma\, r^{-1} = \infty$ implies that $M(f_m) \to 0$ as $m \to \infty$.

Finally, the limit relation (201) can be written in the form $f = f_1 + (f_2 - f_1) + \cdots$ which, by (2), implies that $N(f) \leqq N(f_1) + N(f_2 - f_1) + \cdots$. But the terms of the latter series are non-negative, and its $m$-th partial sum is $M(f_m)$, since $|f_m - f_{m-1}| = f_{m-1} - f_m$, by the inequality (201); cf. (2) and (3). Consequently, $N(f) \leqq M(f_m)$ for every $m$. It follows therefore from $M(f_m) \to 0$ that $N(f) = 0$. Hence it is seen from (2) and (3) that $M(f)$ exists and vanishes

**80.** A curious corollary, which is still elementary, may be formulated as follows:

(ii) *If $d_S(n)$ denotes the number of those divisors of $n$ which are contained in a multiplicative set, $S$, then*

$$\frac{1}{n} \sum_{m=1}^{n} d_S(m) - \sum_{m=1}^{n}{}' \frac{1}{m} \to -\Gamma'(2)|S| \quad \text{as } n \to \infty, \tag{202}$$

*where the accent of the second summation refers to the omission of those summation indices not contained in $S$, and $|S|$ is the asymptotic relative frequency ("density") of $S$.*

In other words, $|S| = M(g)$, if $g(n)$ denotes the characteristic function of $S$. Thus the norm (2) is finite, since $g(n) = O(1)$. Furthermore, $M(g)$ exists, by (i), §79. Hence the assumptions of (VI), §20 are satisfied by the function $f(n)$ for which $f'(n) = g(n)$. Since (15) shows that this $f(n)$ is precisely $d_S(n)$, the proof of (ii) is complete.

It is seen from the formula following ($29_1$) that the determination of the order of the remainder term of the limit relation (202) is Dirichlet's divisor problem, if $S$ is the set of all positive integers.

**81.** In his theory of closed cycles (that is, of multiplicative sets of positive integers), Cesàro ([10], pp. 146–147) has asserted it to be evident that the series $\Sigma\, f'(n)/n$ is convergent whenever $f(n)$ is the characteristic function of a multiplicative set (it is clear from the context that Cesàro's argument is based on (30)-(31), §8). However, if this were really evident, the prime number theorem would be trivial indeed, since it would follow from the definition (22), where $f(n)$ is the characteristic function of the multiplicative set consisting of a single integer, 1. In view of (III) and (IV), it is somewhat surprising that Cesàro's final result turns out to be correct after all:

(iii) *If $f(n)$ is the characteristic function of a multiplicative set, $S$, then the series $\Sigma\, f'(n)/n$ is convergent. Furthermore, $\Sigma\, f'(n)/n = 0$ or $0 < \Sigma\, f'(n)/n \leqq 1$ accord-*

*ing as the sum of the reciprocal values of all primes not contained in $S$ is divergent or convergent; in fact,* (41) *holds in both cases.*

Corollary. *The mean $M(f')$ exists and is* 0 *in both cases.*

In fact, $f'(n) = O(1)$, by (199). Hence the first part of (iii) follows from the first part of (i) in virtue of (X). Correspondingly, the second part of (iii) is implied by (I) and by the second part of (i).

Finally, the assertion of the Corollary follows from (4) and from the first part of (iii). Incidentally, (VI) implies the validity of the converse inference.

**82.** The preceding result contains the answer to a question which, according to oral communications, has repeatedly been raised and often considered (although it does not seem to have been formulated in the literature). The difficulty in answering the question is that, as will be seen in §83, the classical route of the analytic theory of primes, followed after Riemann by Hadamard-de la Vallée-Poussin and Landau-Ikehara, is now blocked by function-theoretical obstacles.

Let $P = Q + R$ be a disjunction of the sequence, $P = \{p\}$, of all primes into two complementary subsequences, $Q = \{q\}$ and $R = \{r\}$. Let $\mu_R(n)$ denote the Möbius function belonging to $R$. By this is meant that $\mu_R(n)$ denotes the multiplicative function for which $\mu_R(p^k)$ is $-1$ or 0 according as neither or either of the conditions $p = r$, $k = 1$ is violated. Accordingly, the generating Dirichlet series is

$$(203)\qquad 1/\zeta_R(s) = \sum \mu_R(n)n^{-s} \quad \text{if} \quad \zeta_R(s) = \prod (1 - r^{-s})^{-1} \equiv \sum (r^*)^{-s}$$

($\sigma > 1$), where the indices $n$, $r$ and $r^*$ respectively run through all positive integers, through the given prime sequence $R = \{r\}$ and through the sequence, $R^* = \{r^*\}$, of those positive integers which are not divisible by any prime, $q$, contained in $Q = P - R$.

Correspondingly, $Q^* = \{q^*\}$ will denote the sequence of the positive integers generated by the prime sequence $Q = \{q\}$. Thus 1 is the only integer common to $R^*$ and $Q^*$. Needless to say, $Q^* + R^*$ is not the set, $P^* = (Q + R)^*$, of all positive integers (unless either $Q = P$ or $R = P$, which is the case of the classical functions $\mu$, $\zeta$). It is also clear that, in the notations of §77, the set $S = Q^*$ is the completely multiplicative set generated by the basis $S^\dagger = Q$. In other words, if $f(n)$ denotes the characteristic function of $S = Q^*$, then

$$(196\text{ bis})\qquad f(p^k) = 1 \quad \text{if} \quad p = q \quad \text{and} \quad f(p^k) = 0 \quad \text{if} \quad p = r,$$

where $k = 1, 2, \cdots$ in both cases. Consequently, not only (198) but also

$$(198\text{ bis})\qquad f'(q^k) = 0, \qquad f'(r^k) = 0, \qquad \text{where } k > 1,$$

follows from (17). This, when compared with the values $\mu_R(p^k)$ assigned before (203), shows that $f'(p^k) = \mu_R(p^k)$ holds for $k = 1, 2, \cdots$ in both cases, $p = q$ and $p = r$. Since $f(n)$ and $\mu_R(n)$ are multiplicative functions, it follows from (39) that $f'(n) = \mu_R(n)$ holds for every positive integer $n$.

Consequently, that particular case of (iii) in which $S$ is *completely* multiplicative, can be formulated as follows:

(iv) *If* $\mu_R(n)$ *denotes the Möbius function belonging to an arbitrary subsequence,* $R = \{r\}$, *of the sequence of all primes, then the series* $\Sigma\, \mu_R(n)/n$ *is convergent. Furthermore,* $\Sigma\, \mu_R(n)/n = 0$ or $0 < \Sigma\, \mu_R(n)/n \leqq 1$ *according as* $\Sigma\, r^{-1} = \infty$ *or* $\Sigma\, r^{-1} < \infty$; *in fact, the reciprocal value of the product* (203) *at* $s = 1$ *represents the value of* $1/\zeta_R(1) = \Sigma\, \mu_R(n)/n$ *in both cases.*

Corollary. *The mean* $M(\mu_R)$ *exists and is* 0 *in both cases.*

Incidentally, the more general theorem, (iii), is readily seen to be implied by (iv).

**83.** The function-theoretical situation referred to at the beginning of §82 can conveniently be described in terms of Ikehara's theorem. This theorem (Wiener [85]; cf. also Karamata [54], Ingham [46], Beurling [4]), when formulated for ordinary Dirichlet series $\Sigma\, f'(n)n^{-s}$ which converge in the half-plane $\sigma > 1$, asserts that, if there exists a constant, $c$, for which the function

$$\phi(s) = \sum f'(n)n^{-s} - c/(s-1), \qquad (\sigma > 1), \tag{204}$$

becomes uniformly continuous in a rectangle $-T < t < T$, $1 < \sigma < 2$, where $s = \sigma + it$, then $M(f')$ exists, and has the value $c$, whenever the following pair of Tauberian conditions is satisfied: $(i)$ $T$ can be chosen arbitrarily large and $(ii)$ $f'(n) = O_L(1)$ as $n \to \infty$.

If $f'(n)$ is given either by (26) or by (22), then $(ii)$ is satisfied, $\Sigma\, f'(n)n^{-s}$ is either $-d \log \zeta(s)/ds$ or $1/\zeta(s)$, and so, if $c = 1$ and $c = 0$ respectively, $(i)$ is reduced to $\zeta(1 + it) \neq 0$, $-\infty < t < \infty$, in both cases, where $\zeta(s)$ is Riemann's zeta-function. Hence, Ikehara's theorem supplies the prime number theorem not only in the form $M(\Lambda) = 1$ but also in the form $M(\mu) = 0$ (usually derived not directly but as a consequence of $M(\Lambda) = 1$). However, it is easy to see that the method fails in the case of §82.

In fact, let $f'(n) = \mu_R(n)$. Then, since $M(\mu_R)$ should vanish, $c$ is 0, and so (204) and (203) imply that $\phi(s) = 1/\zeta_R(s)$. Furthermore, $(ii)$ is satisfied, since $|\, \mu_R(n)\, | \leqq 1$. Finally, $(i)$ requires that $1/\zeta_R(s)$ should go over into a continuous boundary function on the line $\sigma = 1$. Since the latter requirement is satisfied when either $R = P$ or $Q = P$, it is clearly satisfied also in the contiguous cases, represented by prime sequences $\{r\} = R = P - Q$ which are either so "dense" or so "thin" on $P$ that, for instance, $\Sigma\, q^{-1} < \infty$ or $\Sigma\, r^{-1} < \infty$ respectively (in the precise sense of the theory of product measures, the probability of either of such choices of $R$ is zero, since $Q + R = P$ and $\Sigma\, p^{-1} = \infty$).

However, in case of an arbitrary $R$, any continuity in the boundary behavior of $1/\zeta_R(s)$ (or, for that matter, of $\zeta_R(s)$ itself) is doubtful, since the product (203) might acquire a highly pathological natural boundary on the line $\sigma = 1$. It is true that, by the proof of Ikehara's theorem, the assumption of a continuous boundary function can readily be replaced by the assumption that $\phi(s)$ has a

majorized $L$-integrable boundary behavior; that is, by the assumption that there exists for every $T > 0$ an $L$-integrable function $\psi(t)$, $-T < t < T$, satisfying $|\phi(s)| < \psi(t)$ for $-T < t < T$, $1 < \sigma < 2$. But if $R$ is unspecified, the existence of such majorants is about as problematic as that of a continuous boundary function. In fact, very serious difficulties seem to arise if an existence proof for something like an $L$-integrable majorant is attempted not even for every $T > 0$ but merely for some $T = \epsilon$ (or, for that matter, for $t_0 - \epsilon < t < t_0 + \epsilon$, where $t_0$ need not be 0; even if (iv), §82 is granted, all that follows from general theorems is that, for every fixed real $t_0$,

$$1/\zeta_R(s) = o(|s - s_0|^{-1}) \quad \text{as} \quad s \to s_0 \tag{205}$$

holds when $s$ tends to $s_0 = 1 + it_0$ *within* a Stolz wedge).

**84.** Accordingly, it was essential to approach (iv), §82 via (X) or ($IX_1$). Naturally, the proof of the implication, $(L) \to (A)$, of Hardy and Littlewood, an implication used in the proof of ($IX_1$), depends on the prime number theorem itself (that is, on the convergence of $\Sigma\, \mu(n)/n$ in the classical case, $R = P$), and even on somewhat more. In fact, it depends on the convergence of the Dirichlet series of the $l$-th derivative of Riemann's $1/\zeta(s)$ at the point $s = 1$ if $l = 2$, or at any rate if $l = 1 + \epsilon$ (the convergence of this series is known for every $l$, and is needed for $l = 1 + \epsilon$ in order to ascertain the satisfaction of what corresponds to (iii), §10 in case of the linear transformation which sends the Lambertian average into the Abelian one; cf. [37]). However, (iv), §82 appears to be the first instance of a theorem showing that the real Lambertian generator (59) is *actually* superior to the complex Dirichletian generator (25). Thus Wiener's direct proof [85] of the particular (Tauberian) case $f'(n) = O_L(1)$ of the general (Abelian) implication $(L) \to (A)$ of Hardy and Littlewood, a particular case which would have sufficed in the proof of ($IX_1$), acquires a methodical interest.

Needless to say, the difficulties mentioned in §83 go deeper than what corresponds to the non-vanishing of Riemann's $\zeta(s)$ on the line $\sigma = 1$. In fact, the problem involves not only the boundary behavior of $1/\zeta_R(s)$ but that of $\zeta_R(s)$ as well. This is seen by observing that, in a certain respect, $\zeta_R(s)$ and $1/\zeta_R(s)$ are equivalent functions, since $R$ can be replaced by $Q = P - R$ in (203). Actually, the appearance of $\zeta_Q^{\ 2}$ must then introduce some further complications.

**85.** A situation which can more easily be kept under "explicit" control but is delicate enough to illustrate the difficulties involved, results if the class of generalizations, $\mu_R(n)$, of the Möbius function, a class defined before (203), is modified as follows:

For a fixed angle $\theta$, let $\mu_\theta(n)$ denote the multiplicative function attaining for $n = p^k$ the values

$$\mu_\theta(p) = e^{i\theta}, \qquad \mu_\theta(p^2) = \mu_\theta(p^3) = \cdots 0, \tag{206}$$

where $p$ is any prime. For instance, the real-valued functions $\mu_\theta(n)$ are $\mu_\pi(n) = \mu(n)$ and $\mu_0(n) = | \mu(n) |$. The generating Dirichlet series is

$$\sum \mu_\theta(n)n^{-s} = 1/\zeta_\theta(s), \tag{207}$$

if $\zeta_\theta(s)$ is defined by

$$1/\zeta_\theta(s) = \prod_p (1 + e^{i\theta} p^{-s}), \tag{208}$$

where $\sigma > 1$. It is also seen that, if $\omega(n)$ denotes the number of all (or, for that matter, of all distinct) prime divisors of $n$, then

$$\mu_\theta(n) = | \mu(n) | e^{i\theta\omega(n)}, \tag{209}$$

$| \mu(n) |$ being 1 or 0 according as $n$ is or is not square-free.

Let $\sigma = 1$, that is, $s = 1 + it$. Then, since $\Sigma \mid e^{i\theta}/p^s \mid^2 = \Sigma\, p^{-2}$ is a convergent series, (206) shows that (V), §17 is applicable to $g(n) = \mu_\theta(n)n^{-s}$ for every fixed $t$. Accordingly, the series (207) and the product (208) either both converge or both diverge for a fixed $t$, and they represent the same value in case of convergence. Furthermore, the product converges at $s = 1 + it$ if and only if $\Sigma\, e^{i\theta}p^{-s}$, or simply $\Sigma\, p^{-1-it}$, is convergent. On the other hand, Mertens has shown in 1887 (<1896) that $\Sigma\, p^{-1-it}$ is convergent for every $t \neq 0$ at which Riemann's $\zeta(1 + it)$ does not vanish; in fact, the convergence of $\Sigma\, p^{-1-it}$ is uniform on any closed bounded interval consisting of such values $t \neq 0$. Since $\zeta(1 + it) \neq 0$ for every $t$, it follows that the series (207) converges to a continuous function of the real variable $t$ on both open half-lines $0 < | t | < \infty$, $\sigma = 1$. Hence, by a standard extension of Abel's continuity theorem for Dirichlet series, the function $1/\zeta_\theta(s)$, defined by (207) for $\sigma > 1$, is uniformly continuous on both open rectangles $\epsilon < | t | < T$, $1 < \sigma < 2$, where $\epsilon > 0$ and $T > \epsilon$ are arbitrary.

In view of the first item of (IV), §15, the last conclusion could not have been based on (208). On the other hand, the proof had to start with (208), since it was only this start that made possible the reduction of the case of an arbitrary $\theta$ to the classical case, $\theta = \pi$. This warrants the somewhat circuitous nature of the above arrangement of the proof for the existence of a continuous boundary function along both half-lines $0 < | t | < \infty$, $\sigma = 1$.

Now let (207) be identified, for a fixed $\theta$, with the series $\Sigma\, f'(n)n^{-s}$ occurring in (204). Then the assumption (*ii*), made after (204), is satisfied, since $| \mu_\theta(n) | \leqq 1$, by (209). Furthermore, if $c$ in (204) is chosen to be 0, the assumption (*i*), made after (204), is satisfied if and only if the value of $\theta$ is such as to make the function $1/\zeta_\theta(s)$ uniformly continuous on that portion of an immediate vicinity of the point $s = 1$ that is contained in the half-plane $\sigma > 1$. Actually, this requirement of continuity can be replaced by the assumption of an $L$-integrable majorant (in the sense of §83) near $t = 0$. The latter assumption is certainly satisfied if, instead of the case $s_0 = 1$ of (205), an estimate

$$1/\zeta_\theta(s) = o(| s - 1 |^{\epsilon-1}); \qquad s \to 1,\ \sigma > 1,\ (\epsilon = \epsilon_\theta > 0), \tag{210}$$

holds *uniformly* for all directions satisfying $| \arg z | < \frac{1}{2}\pi$, where $z = s - 1$.

Accordingly, if a $\theta$ is such that (210) holds (uniformly in arg $z$) for a sufficiently small positive $\epsilon = \epsilon_\theta$ (or even only such that a corresponding "logarithmic" condition is satisfied when $\epsilon_\theta = 0$), then the mean $M(\mu_\theta)$ of the function $\mu_\theta(n)$ exists and has the assumed value, $c = 0$.

If $\theta/\pi$ is rational, the value of $c$ in (204) in the case $f'(n) = \mu_\theta(n)$ can be calculated from the theory of $L$-series (cf. §56–§57). For instance, $c = M(\mu_\theta)$ is 0 if $\theta = \pi$ but it is $6/\pi^2$ if $\theta = 0$.

# APPENDIX

For the reader's convenience, it seems to be appropriate to append here the proof of the fundamental implication, $(L) \rightarrow (A)$, of Hardy and Littlewood [37]. The notation becomes simpler if their proof is rewritten for the case of Stieltjes integrals. The theorem then states that, if $\alpha(x)$ is a function for which the integrals

$$L(s) = \int_0^\infty \frac{xe^{-sx}}{1 - e^{-sx}}\, d\alpha(x), \qquad A(s) = \int_0^\infty e^{-sx}\, d\alpha(x)$$

exist absolutely* whenever $s > 0$, and if $sL(s)$ tends to a limit, say $C$, as $s \rightarrow 0$, then $sA(s) \rightarrow C$ (cf. §19, where $r = e^{-s}$ in the present notation; the Lambert series (59) and the corresponding power series $\Sigma\, f'(n)r^n$ result by choosing $\alpha(x)$ in $L(s)$ and $A(s)$ to be the step function consisting of the jumps $\alpha(n + 0) - \alpha(n - 0) = f'(n)$, where $n = 1, 2, \cdots$ ).

It can be assumed that $\alpha(x)$ is constant near $x = 0$. It is then obvious that the functions $L(s)$, $A(s)$ and their derivatives tend exponentially to 0, as $s \rightarrow \infty$. Furthermore, it can be assumed without loss of generality that $C = 0$. The assertion then is that $A(s) = o(1/s)$ is implied by $L(s) = o(1/s)$, as $s \rightarrow 0$.

Even if $L(s) = o(1/s)$ is not assumed,

$$L(s) \equiv \int_0^\infty xe^{-sx} \sum_{n=0}^\infty e^{-nsx}\, d\alpha(x) = \sum_{n=1}^\infty \int_0^\infty xe^{-snx}\, d\alpha(x)$$

for every $s > 0$ (the term-by-term integration is justified by the exponential estimate, mentioned before). Since the last integral is identical with $-A'(ns)$, where $A'(s) = dA(s)/ds$,

$$L(s) = -\sum_{n=1}^\infty A'(ns); \quad \text{whence} \quad A'(s) = -\sum_{n=1}^\infty \mu(n)L(ns)$$

follows by Möbius inversion (cf. §25), the legitimacy of this inversion being trivial from the exponential estimates. Accordingly, since $A(\infty) = 0$,

$$A(s) = \int_s^\infty \sum_{n=1}^\infty \mu(n)L(nt)\, dt \equiv \sum_{n=1}^\infty \mu(n) \int_s^\infty L(nt)\, dt.$$

This can be written in the form

$$A(s) = \sum_{n=1}^\infty \frac{\mu(n)}{n} \int_{ns}^\infty L(t)\, dt \equiv \int_0^\infty \left\{ \int_{sx}^\infty L(t)\, dt \right\} d\lambda(x),$$

* The full force of this assumption will not be needed.

if $\lambda(x)$ denotes the $[x]$-th partial sum of the series $\Sigma\, \mu(n)/n$. Since $\lambda(x)$ remains bounded as $x \to \infty$ and vanishes near $x = 0$, partial integration gives

$$A(s) = \int_0^\infty \lambda(x/s)L(x)\, dx,$$

if $x$ is replaced by $x/s$.

The prime number theorem asserts that the series $\Sigma\, \mu(n)/n$ converges to 0, which means that $\lambda(x) = o(1)$ as $x \to \infty$. Actually, it is known that $\lambda(x) = o(1/\log^\alpha x)$ holds not only for $\alpha = 0$ but for *every* $\alpha$, and therefore for *some* $\alpha > 1$ (cf. §84). Since $L(x)$ tends exponentially to 0 as $x \to \infty$, this estimate of $\lambda(x)$ obviously implies that the contribution of the range $1 < x < \infty$ to the last integral is $o(1/s)$ as $s \to 0$.

Hence, in order to prove that $A(s) = o(1/s)$, it is sufficient to show that

$$\int_0^1 \lambda(x/s)L(x)\, dx \equiv \int_0^\epsilon + \int_\epsilon^1 = o(1/s) \quad \text{as} \quad s \to 0.$$

But the truth of this estimate is readily inferred from $\lambda(x) = o(1/\log^\alpha x)$, $x \to \infty$, where $\alpha > 1$, and from the assumption $L(x) = o(1/x)$, $x \to 0$, which was not used thus far.

# BIBLIOGRAPHY


[1] A. Axer, Beitrag zur Kenntnis der zahlentheoretischen Funktionen $\mu(n)$ und $\lambda(n)$, Prace Mat.-Fiz., vol. 21 (1910), pp. 65–95.

[2] A. S. Besicovitch, *Almost Periodic Functions*, Cambridge, 1932.

[3] —— On the density of certain sequences of integers, Math. Annalen, vol. 110 (1934), pp. 336–341.

[4] A. Beurling, Analyse de la loi asymptotique de la distribution des nombres premiers généralisés. I, Acta Math., vol. 68 (1937), pp. 255–291.

[5] V. Brun, Le crible d'Eratosthène et le théorème de Goldbach, Videnskapsselskapets Skrifter (Kristiania), vol. $1920_1$, pp. 1–36.

[6] G. Cantor, Gesammelte Abhandlungen, p. 68, formulae (23) and (25).

[7] —— Vérification jusqu'à 1000 du théorème empirique de Goldbach, C. R. Assoc. Franc., vol. 23 (1894), pp. 117–134 (omitted from Cantor's collected works).

[8] A. Cauchy, Œuvres, ser. 1, vol. 5, pp. 152–159.

[9] E. Cesàro, *Sur diverses questions d'arithmétique*, Bruxelles, 1883.

[10] —— Fonctions énumératrices, Annali di Mat., ser. 2, vol. 14 (1886), pp. 141–158.

[11] J. G. van der Corput, Sur l'hypothèse de Goldbach pour presque tous les nombres pairs, Acta Arithmetica, vol. 2 (1937), pp. 266–290.

[12] H. Davenport and P. Erdös, On sequences of positive integers, Acta Arithmetica, vol. 2 (1936), pp. 147–151.

[13] R. Dedekind, Supplement IX, pp. 385–386 to the 3rd ed. (1879) of Dirichlet's *Vorlesungen über Zahlentheorie*.

[14] —— Commentaries to [27], pp. 292–299 (omitted by the editors of Dedekind's collected works).

[15] —— Gesammelte mathematische Werke, vol. 1 (1930), p. 61.

[16] G. P. Lejeune Dirichlet, Werke, vol. 1, pp. 351–356.

[17] —— ibid., pp. 411–496.

[18] —— ibid., vol. 2, pp. 49–66.

[19] —— ibid., pp. 97–104.

[20] P. Erdös and M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions, Amer. Jour. of Math., vol. 62 (1940), pp. 738–742.

[21] —— and G. Szekeres, Ueber die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem, Acta Litt. Sci. Szeged, vol. 7 (1935), pp. 95–102.

[22] —— and A. Wintner, Additive arithmetical functions and statistical independence, Amer. Jour. of Math., vol. 61 (1939), pp. 713–721.

[23] —— and A. Wintner, Additive functions and almost periodicity ($B^2$), ibid., vol. 62 (1940), pp. 635–645.

[24] L. Euler, Opera Omnia, ser. 1, vol. 8, pp. 284–312.

[25] P. Fatou, Sur les séries entiers à coefficients entiers, Comptes Rendus, vol. 138 (1904), pp. 342–344.

[26] C. F. Gauss, Werke, vol. 1 (ed. 1870), pp. 362–369.

[27] —— ibid., vol. 2, pp. 269–291.

[28] —— ibid., vol. $10_1$, p. 11; for references cf. p. 17.

[29] J. P. Gram, En numerisk funktion, Tidsskrift for Matematik, ser. 5, vol. 2 (1884), pp. 170–181.

[30] —— Undersøgelser angaaende Maengden af Primtal under en given Graense, Skrifter d. K. Danske Videnskabernes Selskab, ser. 6, vol. 2 (1884), pp. 185–308.

[31] —— Studier over nogle numeriske Funktioner, ibid., vol. 7 (1890), pp. 1–34.

[32] —— Tafeln für die Riemannsche Zetafunktion, ibid., ser. 10, vol. 3 (1925), pp. 313–325.

[33] J. Hadamard, Deux théorèmes d'Abel sur la convergence des séries, Acta Math., vol. 27 (1903), pp. 177–184.

[34] G. H. Hardy, Some theorems concerning infinite series, Math. Annalen, vol. 64 (1907), pp. 77–94.

[35] —— Note on Ramanujan's trigonometrical sum $c_q(n)$, Proc. Camb. Phil. Soc., vol. 20 (1921), pp. 263–271.

[36] —— *Ramanujan*, Cambridge, University Press, 1940.

[37] —— and J. E. Littlewood, On a Tauberian theorem for Lambert's series, and some fundamental theorems in the analytic theory of numbers, Proc. Lond. Math. Soc., ser. 2, vol. 19 (1921), pp. 21–29; cf. also ibid., ser. 2, vol. 41 (1936), pp. 257–270.

[38] —— and J. E. Littlewood, Note on Messrs. Shah and Wilson's paper entitled: On an empirical formula connected with Goldbach's Theorem", Proc. Camb. Phil. Soc., vol. 19 (1919), pp. 245–254.

[39] —— and J. E. Littlewood, Some problems of "Partitio numerorum", III, Acta Mathematica, vol. 44 (1923), pp. 1–70.

[40] —— and J. E. Littlewood, Some problems of "Partitio numerorum", V, Proc. Lond. Math. Soc., ser. 2, vol. 22 (1924), pp. 46–56.

[41] —— and S. Ramanujan, [69], pp. 260–275.

[42] P. Hartman and A. Wintner, On the almost periodicity of additive number-theoretical functions, Amer. Jour. of Math., vol. 62 (1940), pp. 753–758.

[43] —— and A. Wintner, Additive functions and almost periodicity, Duke Math. Jour., vol. 9 (1942), pp. 112–119.

[44] ——, E. R. van Kampen and A. Wintner, Asymptotic distributions and statistical independence, Amer. Jour. of Math., vol. 61 (1939), pp. 477–486.

[45] A. E. Ingham, Some asymptotic formulae in the theory of numbers, Jour. Lond. Math. Soc., vol. 2 (1927), pp. 202–208.

[46] —— On Wiener's method in Tauberian theorems, Proc. Lond. Math. Soc., ser. 2, vol. 38 (1933), pp. 459–480.

[47] —— On the "high-indices theorem" of Hardy and Littlewood, Quarterly Jour. of Math., ser. 2, vol. 8 (1937), pp. 1–7.

[48] B. Jessen and A. Wintner, Distribution functions and the Riemann zeta function, Trans. Amer. Math. Soc., vol. 38 (1935), pp. 48–88.

[49] M. Kac, E. R. van Kampen and A. Wintner, Ramanujan sums and almost periodic functions, Amer. Jour. of Math., vol. 62 (1940), pp. 107–114.

[50] E. R. van Kampen, On uniformly almost periodic multiplicative and additive functions, ibid., pp. 627–634.

[51] —— and A. Wintner, On the almost periodic behavior of multiplicative number-theoretical functions, ibid., pp. 613–626.

[52] J. Karamata, Ueber die Hardy-Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes, Math. Zeitschr., vol. 32 (1930), pp. 319–320; also Neuer Beweis und Verallgemeinerung einiger Tauberian-Sätze, ibid., vol. 33 (1931), pp. 295–299.

[53] —— Weiterführung der N. Wienerschen Methode, ibid., vol. 38 (1934), pp. 701–708.

[54] —— Ueber einen Satz von H. Heilbronn und E. Landau, Publ. Math. Univ. Belgrade, vol. 5 (1936), pp. 28–38.

[55] H. D. Kloosterman, On the representation of numbers in the form $ax^2 + \cdots + dt^2$, Acta Math., vol. 49 (1926), pp. 407–464.

[56] L. Kronecker, Quelques remarques sur la détermination des valeurs moyennes, Werke, vol. 5, pp. 303–309.

[57] E. Landau, Ueber die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindestzahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate, Archiv d. Math. u. Phys., ser. 3, vol. 13 (1908), pp. 305–312.

[58] —— Ueber die Aequivalenz zweier Hauptsätze der analytischen Zahlentheorie, Wien, Sitzber., vol. 120 (1911), pp. 973–988.
[59] —— Sur les séries de Lambert, Comptes Rendus, vol. 156 (1913), pp. 1451–1454.
[60] M. LERCH, Zur Theorie der Gaussschen Summen, Math. Annalen, vol. 57 (1903), pp. 554–567.
[61] P. LÉVY, Observations sur la mémoire de M. F. Tricomi, Atti Accad. Sci. Torino, vol. 75 (1939), pp. 177–183.
[62] J. LIOUVILLE, Sur l'expression $\phi(n)$, Jour. de Math., ser. 2, vol. 2 (1857), pp. 110–112.
[63] J. E. LITTLEWOOD, On a class of conditionally convergent infinite products, Proc. Lond. Math. Soc., ser. 2, vol. 8 (1910), pp. 195–199.
[64] A. F. MÖBIUS, Gesammelte Werke, vol. 4, pp. 589–612 and 613–624.
[65] NICHOMACHUS OF GERASA, *Introduction to Arithmetic*, translated by N. L. D'Ooge, Macmillan, 1926.
[66] H. PETERSSON, Ueber die Entwicklungskoeffizienten der automorphen Formen, Acta Math., vol. 58 (1932), pp. 169–215.
[67] A. DE POLIGNAC, Nouvelles recherches sur les nombres premiers, Jour. de Math., ser. 1, vol. 19 (1854), pp. 305–333.
[68] A. RAJCHMAN, Ueber eine paradoxe Eigenschaft gewisser bedingt konvergenter unendlicher Reihen, Math. Zeitschr., vol. 26 (1927), pp. 777–778.
[69] S. RAMANUJAN, Collected Papers, pp. 179–199.
[70] I. SCHUR, Ueber lineare Transformationen in der Theorie der unendlichen Reihen, Crelle Jour., vol. 151 (1920), pp. 79–111.
[71] H. SPÄTH, Ueber Lambertsche Reihen, Math. Zeitschr., vol. 30 (1929), pp. 481–486.
[72] J. J. SYLVESTER, Math. Papers, vol. 2, pp. 709–711.
[73] —— ibid., vol. 4, pp. 84–87.
[74] —— ibid., pp. 88–90.
[75] —— ibid., pp. 734–737.
[76] —— ibid., pp. 738–742.
[77] I. THOMAS, Selections illustrating the history of Greek mathematics, Harvard University Press, vol. 1 (1939), pp. 101–103.
[78] O. TOEPLITZ, Ueber allgemeine lineare Mittelbildungen, Prace Mat.-Fyz., vol. 22 (1913), pp. 113–119.
[79] —— Ein Beispiel zur Theorie der fastperiodischen Funktionen, Math. Annalen, vol. 98 (1928), pp. 281–295.
[80] —— Zur Theorie der Dirichletschen Reihen, Amer. Jour. of Math., vol. 60 (1938), pp. 880–888.
[81] F. TRICOMI, Sulla frequenza dei numeri interi decomponobili nella somma di due potenze $k$-esime, Atti Accad. Sci. Torino 74 (1939), pp. 369–380.
[82] P. TURÁN, Ueber einige Verallgemeinerungen eines Satzes von Hardy und Ramanujan, Jour. Lond. Math. Soc., vol. 11 (1936), pp. 125–133.
[83] L. VINOGRADOW, Some theorems concerning the theory of primes, Recueil Mathématique, vol. 44 (1937), pp. 179–194.
[84] G. N. WATSON, Ueber Ramanujansche Kongruenzeigenschaften der Zerfallungsanzahlen (I), Math. Zeitschr., vol. 39 (1935), pp. 712–731.
[85] N. WIENER, *The Fourier Integral*, Cambridge, 1933, pp. 112–137.
[86] —— AND A. WINTNER, Harmonic analysis and ergodic theory, Amer. Jour. of Math., vol. 63 (1941), pp. 415–426.
[87] —— AND A. WINTNER, On the ergodic dynamics of almost periodic systems, ibid., pp. 794–824.
[88] A. WINTNER, Ueber die statistische Unabhängigkeit der asymptotischen Verteilungsfunktionen inkommensurabler Partialschwingungen, Math. Zeitschr., vol. 36 (1933), pp. 618–629.

[89] —— Ueber die Spektra der Toeplitzschen $D$-Formen, Monatshefte f. Math. u. Phys., vol. 48 (1939), pp. 147–152.

[90] —— Statistics and prime numbers, Nature, vol. 147 (1941), pp. 208–209.

[91] —— On the distribution function of the remainder term of the prime number theorem, Amer. Jour. of Math., vol. 63 (1941), pp. 233–248; cf. ibid., pp. 619–627.

[92] —— On a statistics of the Ramanujan sums, ibid., vol. 64 (1942), pp. 106–114.

[93] —— On the prime number theorem, ibid., pp. 320–326.

[94] —— The distribution of primes, Duke Math. Jour., vol. 9 (1942), pp. 425–430.

[95] —— On an elementary analogue of the Riemann-Mangoldt formula, Bull. Amer. Math. Soc., vol. 48 (1942), pp. 759–762.

# INDEX

(The numbers refer to the pages)